

Gly Gly Arg Arg Xaa Phe
20

(2) INFORMATION FOR SEQ ID NO:130:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Asp Pro Gly Tyr Thr Pro Gly
1 5

(2) INFORMATION FOR SEQ ID NO:131:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: /note= "The Second Residue Can Be Either a Pro or Thr"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

Xaa Xaa Gly Phe Thr Gly Pro Gln Phe Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:132:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 9 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: /note= "The Third Residue Can Be Either a Gln or Leu"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Xaa Pro Xaa Val Thr Ala Tyr Ala Gly
1 5

(2) INFORMATION FOR SEQ ID NO:133:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

Xaa Xaa Xaa Glu Lys Pro Phe Leu Arg
1 5

(2) INFORMATION FOR SEQ ID NO:134:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

Xaa Asp Ser Glu Lys Ser Ala Thr Ile Lys Val Thr Asp Ala Ser
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:135:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Ala Gly Asp Thr Xaa Ile Tyr Ile Val Gly Asn Leu Thr Ala Asp
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:136:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136;

Ala Pro Glu Ser Gly Ala Gly Leu Gly Gly Thr Val Gln Ala Gly
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

Xaa Tyr Ile Ala Tyr Xaa Thr Thr Ala Gly Ile Val Pro Gly Lys Ile
1 5 10 15

Asn Val His Leu Val
20

(2) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 882 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

| | |
|--|-----|
| GGAACGCTGT CCGGCCCTTT GCGGTGATCG GTTTCGCCCTC GCTGGCGGTC GCGGTGGCCG | 60 |
| TCACCATCG ACCGACCGCG GCCTCAAAC CGGTAGAGG ACACCAAAAC GCGCAGCCAG | 120 |
| GGAAGTTCAT GCGTTGTTC CCGACGCAAC AGCAGGCGCC GGTCCCGCCG CCTCCGCCCG | 180 |
| ATGATCCAC CCGTGGATTC CAGGCGGSCA CCATTCCGGC TGTACAGAAC GTGGTGCCGC | 240 |
| GGCGGGGTAC CTCACCGGG GTGGGTGGGA CGCCCGCTTC GCGTGGCCG GAGCGGCGG | 300 |
| CGGTGCCCGG TGTGTGCGT GCGCCGGTGC CATTCCCGGT CCGGATCATC ATTCCCGCGT | 360 |
| TCCCGGCTTG GCAGCCTGGA ATCCSACCA TCCCLACGCG ACCCGCGACG ACCCGGGTGA | 420 |
| CCACGTCCGC GAGGAGGCG CCGACCTGCG CCGCGACCA GCGGGTGACC ACCCGGCCAA | 480 |
| CGACGCGCGC GACCAAGCG GTGACCAAG CCGCAACGAC GCGCGGACG ACGCGCGTGA | 540 |
| CCACGCGCAC AACGACCGTC GCGCCGACGA CGGTGCGCG GACGACCGTC GCTCCGACCA | 600 |
| CGTCCGCGCG GACCAAGGTC GCTCCAGCA CGGCGACGCG GACGACCGTC GCTCCGACCG | 660 |

147

```

CGACGCAGCA GCCCAGGCAA CAACCAACCC AACAGATGCC AACCCAGCAG CAGACCCCTGG      720
CCCCGCAGAC GGTGGGCCCC GCTCCGCAGC CCCCCTCCGG TGGCCGCAGC GGCAGCSCCG      780
GGGGCGACTT ATTCCGCGCG TTCTGATCAC GTCCGCGGCT TCACTAGCGT CCGAGGACAT      840
GGCCGCTGAT GCCTGACGCG TCGTGTCTCC CTGTCTCAAC GA                      882

```

(2) INFORMATION FOR SEQ ID NO:139:

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 815 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

```

(ii) MOLECULE TYPE: DNA (genomic)

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:139:

```

CCATCAACCA ACCGCTCGCG CCGCCCGGCG CCGCGGATCG GCGGTGGGCG CCACGCGGCG      60
CGGTGCTTCC GGTGCCCCCG TTGCCGCGGT CGCGGCGGTC GCGCGGAGC GGTGCGTGC      120
CTAGGGGCGCT GTTACCGCCC TGGTTGGCGG GACGCGCGCC GGCACGACCG GTACCGCGCA      180
TGGCGCGGCTT GCGCGCGGCG GCACGCTTGC CACCGTTGCC ACCGTTGCCA CGGTGCGGCA      240
CCAGCGACGC GCGCGGACCA CCGCGCGGCG GCGCGCGGCG GCGACGCGCG GCGTACCGGT      300
TCGTGCGGCT ACCGCGCGCA CCGCGGTTGC CCGCGTCACC GCGACGCGAA CTACCGCGCG      360
ACGCGGCGTG CCGCGCGGCG CCGCGCGGCG CCGCATTCGC ACCGCGGTCG CCGCGGCGTG      420
GGAGTGGCGC GATTAGGGCA CTGACCGGCG CAACGAGCGC AAGTACTGTC GTTCACCGAG      480
CAGTTCCAGA CGACACGACA GCACGCGGCT GTCGCGGAGC TGGGTGAAT GGCACCGGAT      540
AGCGGCTAGC TGTGCGCTGC GTTCACCGTC GATCATGATG TCGAGGTGAC GTTGACCGCG      600
CCCCCGGAAG GAGCGGCTGA ACTCGCGGCT GCGCGGATCG GCGATCGGTT GCGGCGGTC      660
CCAGGCGCAT ACCGCGGATG CCGGTGTCGA AGCGCGGCGG AGCGGAGCTT CGGTGCGGCG      720
ACGCTGGTCC GGTGCGGCTG TTACCGCGCT GTCTCGAAGC ACGAGTAGCA GTCTGCTCC      780
GGCGAGGCGA TCCACGACCG GTTGGGTCAG CTCGT                      815

```

(2) INFORMATION FOR SEQ ID NO:140:

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1152 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

```

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

```

ACCAGCCSCC GGCTGAGGTC TCAGATCAGA GAGTCTCCGG ACTCACCAGG GCGGTTGAGC      60
CITCTCCGAG AACAACTGCT GAGGATCTTC GCCCGCGAAA CAGCGGCTGA ITTGAGCGCTC      120
TATGACCGGT TGAACGACGA GATCATCCGG CAGATTGATA TGCCACCGGT GGGCTAACAG      180
GTGCGCAAGA TGGTGACGCT GTATGTCTCG GACTCCGTCT CCGGATCAG CTTTGCCGAC      240
GGCCGGGTGA TCGTGTGGAG CGAGGAGCTC GCGGAGAGCC AATATCCGAT CGAGACGCTG      300
GACGGCATCA CGCTGTTTGG GCGGCGGAGC ATGACAACGC CCTTCATCGT TGAGATGCTC      360
AAGCGTGAGC GCGACATCCA GCTCTTCAGC ACCGACGGCC ACTACCAGGG CCGGATCTCA      420
ACACCGGAGC TGTATACGCG GCGCGGCTC CGTCAGCAAG TTCACCGCAC CGACGATCTT      480
GCGTCTCGCC TGTCTTAAAG CAAGCGGATC GTGTGAGGGA AGATCTGAA TCACGAGGCC      540
TTGATTCGGG CACACAGCTC GGGGCAAGAC GTTGTGAGA GCATCCGCAC GATGAAGCAC      600
TCGCTGGCCT GCGTCGATCG ATCGGCTCC CTGGCGGAGT TGAACGGTCT CAGGGCAAT      660
GCCGCAAGG CATACCTCAC CGCGCTGGGG CATCTCGTCC CGCAGGAGTT CGCATTCAG      720
GGGCTCTGGA CTCGGGCGCC GTTGAGCGCC TTCAACTCGA TGGTCAGCCT CCGCTATTCG      780
CTGCTGTACA AGAACATCAT AGGGGCGATC GAGGCTCACA GCCTGAACGC GTATATCGGT      840
TTCCTACACC AGGATTCAGG AGGGCAGGCA AGGTCTCGTG CCGAATTCGG CACGAGCTCC      900
GCTGAAACCG CTGGCGGGCT GCTCAGTGCC CTTACGTAAT CCGCTGCGCC CAGGCGGGCC      960
CGCCGGCCGA ATACCAGGAG ATCGGACAGC GAATTCGGGC CCAGCGGGTT GGAGCCGTGC      1020
ATACCGCCGG CACACTCACC GGCAGCGAAG AGGCTGSCA CCGTGGCGGC GCGGCTGTCC      1080
GCGTCTACTT CGACACCGCC CATCACGTAG TGACAGGTCG GCGGACTTC CATTGCTCTC      1140
GTTCGGCAGG AG                                     1152

```

(2) INFORMATION FOR SEQ ID NO:141:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 655 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

```

CTCGTCCCGA TTCGGCAGGG TGTAATGCC GGTGGTGTAT GCGGCTGAG TGCCGACGAC      60
CAGCAATGCG GCAACAGCAC GATCCCGGT CAACGACGCC ACCCGGTCCA CGTGGGGGAT      120

```

```

CCGCTCGAGT CCGCCTTGGG CCGCTCTTTC CTTGGGCAGG GTCATCCGAC GTGTTTCCGC      180
CCGCGTTTGC CCGCATTATG CCGGCGCGCC GCGTCGGCGG GCGGTATGCG CGAANGTCG      240
ATCAGCACAC CCGAGATACG GGTCTGTCCA AGCTTTTGA GCGTCGCGCG GGGCAGCTTC      300
CGCGGCAATT CTACTAGCGA GAAGTCTGGC CCGATACGGA TCTGACGAA GTGCTGCGG      360
TGCAGGECAC CTCATTGGC GATGGCGCGG AGGATGGCGG CTGGAGCGAT CTTGTGCGGC      420
TTGCGGACGG CGACGCGTA GGTGGTCAAG TCCGCTCTAC GCTTGGGCTT TTGCGGACGG      480
TCCGAGCGCT GGTGCGGTT CCGCGCGCAA AGCGCGCGGT CCGGTGCCAT CAGGAATGCC      540
TCACGCGCGC GGCAGTGGC GCGCACTGGC GCGGCGATGT CAGCATCGG GACATCATGC      600
TGGGTTTCAT ACTCTCGAC CAGTCGGCGG AACAGCTCGA TTCCGCGACC GCCCA      655

```

(2) INFORMATION FOR SEQ ID NO:142:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 267 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:142:

```

Asn Ala Val Val Ala Phe Ala Val Ile Gly Phe Ala Ser Leu Ala Val      1
1      5      10
Ala Val Ala Val Thr Ile Arg Pro Thr Ala Ala Ser Lys Pro Val Glu      20
20     25     30
Gly His Gln Asn Ala Gln Pro Gly Lys Phe Met Pro Leu Leu Pro Thr      35
35     40     45
Gln Gln Gln Ala Pro Val Pro Pro Pro Pro Asp Asp Pro Thr Ala      50
50     55     60
Gly Phe Gln Gly Gly Thr Ile Pro Ala Val Gln Asn Val Val Pro Arg      65
65     70     75     80
Pro Gly Thr Ser Pro Gly Val Gly Gly Thr Pro Ala Ser Pro Ala Pro      85
85     90     95
Gln Ala Pro Ala Val Pro Gly Val Val Pro Ala Pro Val Pro Ile Pro      100
100    105    110
Val Pro Ile Ile Ile Pro Pro Phe Pro Gly Trp Gln Pro Gly Met Pro      115
115    120    125
Thr Ile Pro Thr Ala Pro Pro Thr Thr Pro Val Thr Thr Ser Ala Thr      130
130    135    140
Thr Pro Pro Thr Thr Pro Pro Thr Thr Pro Val Thr Thr Pro Pro Thr

```

150

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Thr | Pro | Pro | Thr | Thr | Pro | Val | Thr | Thr | Pro | Pro | Thr | Thr | Pro | Pro | Thr |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Thr | Pro | Val | Thr | Thr | Pro | Pro | Thr | Thr | Val | Ala | Pro | Thr | Thr | Val | Ala |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Pro | Thr | Thr | Val | Ala | Pro | Thr | Thr | Val | Ala | Pro | Thr | Thr | Val | Ala | Pro |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Ala | Thr | Ala | Thr | Pro | Thr | Thr | Val | Ala | Pro | Gln | Pro | Thr | Gln | Gln | Pro |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Thr | Gln | Gln | Pro | Thr | Gln | Gln | Met | Pro | Thr | Gln | Gln | Gln | Thr | Val | Ala |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Pro | Gln | Thr | Val | Ala | Pro | Ala | Pro | Gln | Pro | Pro | Ser | Gly | Gly | Arg | Asn |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Gly | Ser | Gly | Gly | Gly | Asp | Leu | Phe | Gly | Gly | Phe | | | | | |
| | | | 260 | | | | | 265 | | | | | | | |

(2) INFORMATION FOR GEO ID NO: 143:

(1) SEQUENCE CHARACTERISTICS:

(A) Length: 174 amino acids

(E) TYPE: amino acid

(C) STRANDEDNESS: single

(10) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

{xi} SEQUENCE DESCRIPTION: SEQ ID NO:143.

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 115 | Asn | Gln | Pro | Leu | Ala | Pro | Pro | Ala | Pro | Asp | Pro | Pro | Ser | Pro |
| 1 | | | | 5 | | | | | | 10 | | | | 15 |
| Pro | Arg | Pro | Pro | Val | Pro | Pro | Val | Pro | Pro | Leu | Pro | Pro | Ser | Pro |
| | | | 20 | | | | | 25 | | | | | 30 | |
| Ser | Pro | Pro | Thr | Gly | Trp | Val | Pro | Arg | Ala | Leu | Leu | Pro | Pro | Trp |
| | | | 35 | | | 40 | | | | | | 45 | | Leu |
| Ala | Gly | Thr | Pro | Pro | Ala | Pro | Pro | Val | Pro | Pro | Met | Ala | Pro | Leu |
| | 50 | | | | | 55 | | | | | 60 | | | Pro |
| Pro | Ala | Ala | Pro | Leu | Pro | Pro | Leu | Pro | Pro | Leu | Pro | Pro | Leu | Pro |
| 65 | | | | | | 70 | | | | 75 | | | | 80 |
| Ser | His | Pro | Pro | Arg | Pro | Pro | Ala | Pro | Pro | Ala | Pro | Pro | Ala | Pro |
| | | | | 85 | | | | | 90 | | | | | 95 |
| Ala | Cys | Pro | Phe | Val | Pro | Val | Pro | Pro | Ala | Pro | Pro | Leu | Pro | Ser |
| | | | 100 | | | | | 105 | | | | | 110 | |
| Pro | Pro | Thr | Glu | Leu | Pro | Ala | Asp | Ala | Ala | Cys | Pro | Pro | Ala | Pro |
| | | | 115 | | | | 120 | | | | | 125 | | |

151

Ala Pro Pro Leu Ala Pro Pro Ser Pro Pro Ala Gly Ser Ala Ala Ile
 130 135 140

Arg Ala Leu Thr Gly Ala Thr Ser Ala Ser Thr Leu Gly His Arg Ala
 145 150 155 160

Leu Pro Asp Asp Thr Thr Ala Arg Gly Cys Arg Arg Thr Gly
 165 170

(2) INFORMATION FOR SEQ ID NO:144:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

Gln Pro Pro Ala Glu Val Ser Asp Gln Arg Val Ser Gly Leu Thr Gly
 1 5 10 15

Ala Val Gln Pro Ser Pro Arg Thr Thr Ala Glu Asp Pro Arg Pro Arg
 20 25 30

Asn Arg Arg
 35

(2) INFORMATION FOR SEQ ID NO:145:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 104 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

Arg Ala Asp Ser Ala Gly Cys Thr Cys Arg Trp Cys Xaa Pro His Glu
 1 5 10 15

Cys Arg Arg Pro Ala Met Arg Gln Gln His Gly Ser Arg Ser Thr Thr
 20 25 30

Pro Pro Gly Pro Arg Gly Arg Ser Ala Arg Val Arg Pro Gly Arg Leu
 35 40 45

Phe Pro Trp Ala Gly Ser Ser Asp Val Phe Pro Pro Trp Phe Ala Ala
 50 55 60

Ile Met Pro Ala Arg Arg Val Gly Arg Pro Val Trp Pro Xaa Val Asp
 65 70 75 80

152

Gln His Thr Arg Asp Thr Gly Leu Cys Lys Leu Phe Gln Arg Arg Ala
 85 95 98

Gly Gln Leu Arg Arg Gln Phe Tyr
 100

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

GGATCATAT GGGCATCAT CATCATCATC ACGTGATCGA CATCATCGGG ACC

53

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR Primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

CCTGAATTCA GGCCTCGGTT GCGCCGCGCT CATCTGAAC GA

42

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR Primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:148:

GGATCCTGCA GGCTCGAAAC CACCGACCGG T

31

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CTCTGAATTC AGCGCTGGAA ATCGTCCCGA T

31

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:151:

GCATCCAGCG CTGAGATCAA GACCGATGCG CCG

33

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:151:

154

GAGAGAAATTC TCAGAAGCCC ATTTGCCGAGG ACA

33

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1993 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 152..1273

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:152:

| | |
|--|-----|
| TGTTCTTCTCA CCCCAGGCTG GTGGAGGAAG GGGCCACCGA ACAGCTGTTT TCCTCCCGCA | 60 |
| AGCATGCGGA AACCCGCCGA TACGTGCGCG GACTGTCCGG GGACGTCAAG GAGCCCAAGC | 120 |
| GCGGAAATTC AAGAGCACAG AAGGTATGG C GTG AAA ATT CGT TTG CAT ACG | 172 |
| Val Lys Ile Arg Leu His Thr | |
| 1 5 | |
| CTG TTG GCC GTG TTG ACC GCT GCG CCG CTG CTG CTA GCA GCG GCG GCG | 220 |
| Leu Leu Ala Val Leu Thr Ala Ala Pro Leu Leu Leu Ala Ala Ala Gly | |
| 10 15 20 | |
| TGT GCG TCG AAA CCA CCG AGC GGT TCG CCT GAA ACG GGC GCG GCG GCG | 260 |
| Cys Gly Ser Lys Pro Pro Ser Gly Ser Pro Glu Thr Gly Ala Gly Ala | |
| 25 30 35 | |
| GGT ACT GTC GCG ACT ACC CCC GCG TCG TCG CCG GTG ACG TTG GCG GAG | 316 |
| Gly Thr Val Ala Thr Thr Pro Ala Ser Ser Pro Val Thr Leu Ala Glu | |
| 40 45 50 55 | |
| ACC GGT AGC ACG CTG CTC TAC CCG CTG TTC AAC CTG TCG GGT CCG GCC | 364 |
| Thr Gly Ser Thr Leu Leu Tyr Pro Leu Phe Asn Leu Trp Gly Pro Ala | |
| 60 65 70 | |
| TTT CAC GAG AGG TAT CCG AAC GTC ACG ATC ACC GCT CAG GCG ACC GGT | 412 |
| Phe His Glu Arg Tyr Pro Asn Val Thr Ile Thr Ala Gln Gly Thr Gly | |
| 75 80 85 | |
| TCT GGT GCC GGG ATC GCG CAG GCG GCG GCG GCG ACG GTC AAC ATT GCG | 460 |
| Ser Gly Ala Gly Ile Ala Gln Ala Ala Ala Gly Thr Val Asn Ile Gly | |
| 90 95 100 | |
| GCC TCC GAC GCC TAT CTG TCG GAA GGT GAT ATG GCC GCG CAC AAG GGG | 508 |
| Ala Ser Asp Ala Tyr Leu Ser Glu Gly Asp Met Ala Ala His Lys Gly | |
| 105 110 115 | |
| CTG ATG AAC ATC GCG CTA GCC ATC TCC GGT CAG CAG GTC AAC TAC AAC | 556 |
| Leu Met Asn Ile Ala Leu Ala Ile Ser Ala Gln Gln Val Asn Tyr Asn | |
| 120 125 130 135 | |

155

| | | |
|-------------------------------------|----------------------------------|------|
| CTG CCC GGA GTG AGC GAG CAC CTC AAG | CTG AAC GGA AAA GTC CTG GCG | 604 |
| Leu Pro Gly Val Ser Gln His Leu Lys | Leu Asn Gly Lys Val Leu Ala | |
| 140 | 145 150 | |
| GCC ATG TAC CAG GGC ACC ATC AAA ACC | TGG GAC GAC CCG CAG ATC GCT | 652 |
| Ala Met Tyr Gln Gly Thr Ile Lys Thr | Trp Asp Asp Pro Gln Ile Ala | |
| 155 | 160 165 | |
| GCG CTC AAC CCC GGC GTG AAC CTG CCC | GGC ACC GCG GTA GTT CCG CTG | 700 |
| Ala Leu Asn Pro Gly Val Asn Leu Pro | Gly Thr Ala Val Val Pro Leu | |
| 170 | 175 180 | |
| CAC CGC TCC GAC GGG TCC GGT GAC ACC | TTC TTG TTC ACC CAG TAC CTG | 748 |
| His Arg Ser Asp Gly Ser Gly Asp Thr | Phe Leu Phe Thr Gln Tyr Leu | |
| 185 | 190 195 | |
| TCC AAG CAA GAT CCC GAG GGC TGG GGC | AAG TCG CCC GGC TTC GGC ACC | 796 |
| Ser Lys Gln Asp Pro Gln Gly Trp Gly | Lys Ser Pro Gly Phe Gly Thr | |
| 200 | 205 210 215 | |
| ACC GTC GAC TTC CCG GCG GTG CCG GGT | GCG CTG GGT GAG AAC GGC AAC | 844 |
| Thr Val Asp Phe Pro Ala Val Pro Gly | Ala Leu Gly Gln Asn Gly Asn | |
| 220 | 225 230 | |
| GGC GGC ATG GTG ACC GGT TGC GCC GAG | ACA CCG GGC TGC GTG GCC TAT | 892 |
| Gly Gly Met Val Thr Gly Cys Ala Gln | Thr Pro Gly Cys Val Ala Tyr | |
| 235 | 240 245 | |
| ATC GGC ATC AGC TTC CTC GAC CAG GGC | AGT CAA CCG GGA CTC GGC GAG | 940 |
| Ile Gly Ile Ser Phe Leu Asp Gln Ala | Ser Gln Arg Gly Leu Gly Gln | |
| 250 | 255 260 | |
| GCC CAA CTA GGC AAT AGC TCT GGC AAT | TTC TTG TTG CCC GAC GCG CAA | 988 |
| Ala Gln Leu Gly Asn Ser Ser Gly Asn | Phe Leu Leu Pro Asp Ala Gln | |
| 265 | 270 275 | |
| AGC ATT CAG GCC GCG GCG GCT GGC TTC | GCA TCG AAA ACC CCG GCG AAC | 1036 |
| Ser Ile Gln Ala Ala Ala Ala Gly Phe | Ala Ser Lys Thr Pro Ala Asn | |
| 280 | 285 290 295 | |
| CAG GCG ATT TCG ATG ATC GAC GGG CCC | GCC CCG GAC GGC TAC CCG ATC | 1084 |
| Gln Ala Ile Ser Met Ile Asp Gly Pro | Ala Pro Asp Gly Tyr Pro Ile | |
| 300 | 305 310 | |
| ATC AAC TAC GAG TAC GCC ATC GTC AAC | AAC CCG CAA AAG GAC GCC GCC | 1132 |
| Ile Asn Tyr Gln Tyr Ala Ile Val Asn | Asn Arg Gln Lys Asp Ala Ala | |
| 315 | 320 325 | |
| ACC GCG CAG ACC TTG CAG GCA TTT CTG | CAC TGG GCG ATC ACC GAC GGC | 1180 |
| Thr Ala Gln Thr Leu Gln Ala Phe Leu | His Trp Ala Ile Thr Asp Gly | |
| 330 | 335 340 | |
| AAC AAG GCC TCG TTC CTC GAC CAG GTT | CAT TTC CAG CCG CTG CCG CCC | 1228 |
| Asn Lys Ala Ser Phe Leu Asp Gln Val | His Phe Gln Pro Leu Pro Pro | |
| 345 | 350 355 | |
| GCG GTG GTG AAG TTG TCT GAC GCG TTG | ATC GCG ACC ATT TCC AGC | 1276 |
| Ala Val Val Lys Leu Ser Asp Ala Leu | Ile Ala Thr Ile Ser Ser | |
| 360 | 365 370 | |
| TAGCCTCGTT GACCACCAGC CGACAGCAAC | CTCCGTCGGG CCATCGGGCT GCTTTGCGGA | 1333 |

```

GCATGCTGGC CCGTCCCGGT GAAGTGGGCC GCGGTGGCCC GGCATCCGG TGGTTGGGTC 1393
GGATAGGTGC GTTGATCCCG CTGCTTGGGC TGGTCTTGGT GCTGGTGGTC CTGGTCATCG 1453
AGGCGATGGG TGGATCAGG CTCACCGGCT TGCATTTCTT CACCGCCACC GAATGGAAATC 1513
CAGGCAACAC CTACGGCGAA ACCGTTGTCA CCGAGCCSTC GCCCATCCGG TCGCCGCCCTA 1573
CTACGGGGCG TTGCGGCTGA TCGTCGGGAC GCTGGCGACC TCGGCAATCG CCGTGATCAT 1633
CGGGGTGCCC GTCTCTGTAG GAGCGGCGCT GGTGATCGTC GAACGGGTGC CGAAACGGTT 1693
GGCCGAGGCT GTGGGAATAG TCCTGGAAAT GCTCGCCGGA ATCCCCAGCG TGCTGCTCCG 1753
TTTGIGGGGG GCAATGACCT TCGGGCCGTT CATCGCTCAT CACATCGCTC CCGTGATCGC 1813
TCACAACGCT CCGATGTGC CCGTGTGAA CTACTTGGCC GCGACCCCG GCACCGGGGA 1873
GGCATGTTG GTGTCGGTC TGGTGTGGC GGTGATGGTC GTTCCATTA TCGCCACCAC 1933
CACTCATGAC CTGTTCCGSC AGTGCCGCT GTGCCCCCG GAGGGCGGGA TCGGAATTC 1993

```

(2) INFORMATION FOR SEQ ID NO:153:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 374 amino acids

(B) TYPE: amino acid

(C) TOPOLOGY: linear

(11) MOLECULE TYPE: protein

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:153:

```

Val Lys Ile Arg Leu His Thr Leu Leu Ala Val Leu Thr Ala Ala Pro
 1             5             10             15
Leu Leu Leu Ala Ala Ala Gly Cys Gly Ser Lys Pro Pro Ser Gly Ser
 20             25             30
Pro Glu Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro Ala Ser
 35             40             45
Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr Pro Leu
 50             55             60
Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn Val Thr
 65             70             75             80
Ile Thr Ala Gln Gly Thr Gly Ser Gly Ala Gly Ile Ala Gln Ala Ala
 85             90             95
Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser Glu Gly
100             105             110
Asp Met Ala Ala His Lys Gly Leu Met Asn Ile Ala Leu Ala Ile Ser
115             120             125
Ala Gln Gln Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His Leu Lys
130             135             140

```

157

Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Gln Gly Thr Ile Lys Thr
 145 150 155 160
 Trp Asp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gly Val Asn Leu Pro
 165 170 175
 Gly Thr Ala Val Val Pro Leu His Arg Ser Asp Gly Ser Gly Asp Thr
 180 185 190
 Phe Leu Phe Thr Gln Tyr Leu Ser Lys Gln Asp Pro Glu Gly Trp Gly
 195 200 205
 Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val Pro Gly
 210 215 220
 Ala Leu Gly Glu Asn Gly Asn Gly Gly Met Val Thr Gly Cys Ala Glu
 225 230 235 240
 Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp Gln Ala
 245 250 255
 Ser Gln Arg Gly Leu Gly Glu Ala Gln Leu Gly Asn Ser Ser Gly Asn
 260 265 270
 Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Gly Phe
 275 280 285
 Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro
 290 295 300
 Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile Val Asn
 305 310 315 320
 Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu
 325 330 335
 His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp Gln Val
 340 345 350
 His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu
 355 360 365
 Ile Ala Thr Ile Ser Ser
 370

(2) INFORMATION FOR SEQ ID NO:154:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1993 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(3) SEQUENCE DESCRIPTION: SEQ ID NO:154:

TGTCTCTCGA CGGCAGGCTG GTGGAGGAAG GGCCACCGA ACAGCTGTC TCCTCCCGA

60

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| AGCATGCCGA | AACCCCCCGA | TACGTCGCCG | GACTGTCGCG | GGACGTCAG | GACCCCAAGC | 120 |
| GCGGAAATTC | AAGAGCCACG | AAAGGTATGG | CGTGAAAATT | CGTTTCATA | CGCTGTTGGC | 180 |
| CGTGTGACC | GCTGGGCGGC | TGCTGCTAGC | AGCGGCGGGC | TGTGGCTCGA | AACCACCGAG | 240 |
| CGGTTCCCTT | GAAACGGGCG | CCGGCCTCGG | TACTGTCGCG | ACTACCCCGG | CGTCCTCGCC | 300 |
| GCTGACGTTG | GCGGAGACCG | GTAGCAGCGT | GCTCTACCCG | CTGTTCAACC | TGTGGGCTCC | 360 |
| GGCCTTTCAC | GAGAGGTATC | CGAACGTCAC | GATCACCGCT | CAGGSCACCG | GTTCTGGTGC | 420 |
| CGGGATCGCG | CAGGCGGCGG | CCGGGACGGT | CAACATTGGG | GCCTCCGACG | CCATCTCTGC | 480 |
| GGAAGGTGAT | ATGCCCCGCG | ACRAGGGGCT | GATGAACATC | GCGGTAGCCA | TCTCCGCTCA | 540 |
| GCAGGTCAAC | TACAACTGCT | CCGGAGTGAG | CGAGCACCTC | AAGTGAACG | GAAAAGTGGT | 600 |
| GGCGGCCATG | TACCAGGGCA | CCATCAAAAC | CTGGGACGAC | CCGCAGATCG | CTGGGCTCAA | 660 |
| CCCCGGCGTG | AACCTGCCCC | GCACCGCGGT | ACTTCGCTCG | CACCGCTCCG | ACGGGTCCGG | 720 |
| TGACACCTTC | TTGTTACGCG | AGTACCTGTC | CAGCGAAGAT | CCCGAGGGCT | GGGGCAAGTC | 780 |
| CCCCGGCTTC | GGCACCACCG | TGCACTTCGG | GGCGGTGCGG | GCTGCGCTCG | GTGAGAACCG | 840 |
| CAACGGCGGC | ATGCTGACCG | GTTCGCGCGA | GACACCGGGC | TGCTGGGCTT | ATATCGGCAT | 900 |
| CAGCTTCCTC | GACCAGGCCA | GTCAACGGGG | ACTCGGCGAG | GCCCAACTAG | GCAATAGCTC | 960 |
| TGGCAATTTC | TTGTTGCCCC | ACGCGCAAGG | CATTGAGGGC | GCGGGCGCTE | GCTTGGCATC | 1020 |
| GAAACCCCGG | GCGAACCCAG | CGATTTGAT | GATCGACGGG | CCCCCCCCCG | ACGGCTACCC | 1080 |
| GATCATCAAC | TACGAGTACG | CCATCGTCAA | CAACCGGCAA | AAGGACGCGG | CCACCGCGCA | 1140 |
| GACCTTCAG | GCAATTCTCG | ACTGGGCGAT | CACCGACGGC | AACAAGGCTT | CGTTCTCTGA | 1200 |
| CCAGGTTGAT | TTCCAGCCCG | TGCGGCGCGC | GGTGGTGAAG | TTGTCGACG | GCTTGATCGC | 1260 |
| GACGATTTCC | AGCTAGGCTC | GTGACCACCG | ACGCGACAGC | AACCTCCCTC | GGGCCATCGG | 1320 |
| GCTGCTTTGC | GGAGCATGCT | GCCCCGTGCG | GGTGAAGTGG | GCGGCGCTCG | CCCGGCCATC | 1380 |
| CGGTGGTTGG | GTGGGATAGG | TGCGGTGATC | CCGCTGCTTG | CGCTGGTCTT | GTTGCTGGTG | 1440 |
| GTGCTGGTCA | TGGAGGCGAT | GGTGGCGATC | AGGCTCAAGG | GTTTGCATTT | CTTCACCGCC | 1500 |
| ACCGAATGGA | ATCCAGGCAA | CACCTACGGC | GAAACCGTTG | TCACCGACCG | GTGCCCCATC | 1560 |
| CGGTGGGCGC | CTACTACGGG | GCTTTCGCGC | TGATGCTCGG | GACGCTGGCG | ACCTCGGCGA | 1620 |
| TGCGCTGAT | CATCGCGGTG | CCGCTCTCTG | TAGGAGCGGC | GCTGGTGATC | GTGGAACGCG | 1680 |
| TGCTGAAGCG | GTGCGCGGAG | GCTGTGGGAA | TAGTCTCTGA | ATTGCTCGCC | GGAATCCCGA | 1740 |
| GCTGGGTGCT | CGGTTTGTGG | GGGCGCATGA | CGTTGCGGCG | GTTTATCGCT | CATCACATCG | 1800 |
| CTCGGTTGAT | CGCTACAAAC | GCTCCCGATG | TGCGGGTGGT | GAACTACTTG | CGCGGCGAGC | 1860 |
| CGGCGACCGG | GGAGGGCATG | TGGTGTGCGG | GTCTGGTGGT | GCGGGTGATG | GTGCTTCCCA | 1920 |

TTATCGCCAC CACCACTCAT GACCTGTTCC GCGAGGTGCC GGTGTTGCCC CCGGAGGGCC 1980
CGATCGGGA TTC 1993

(2) INFORMATION FOR SEQ ID NO:155:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(a1) SEQUENCE DESCRIPTION: SEQ ID NO:155:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Ile | Arg | Leu | His | Thr | Leu | Leu | Ala | Val | Leu | Thr | Ala | Ala | Pro | 1 | 5 | 10 | 15 |
| Leu | Leu | Leu | Ala | Ala | Ala | Gly | Cys | Gly | Ser | Lys | Pro | Pro | Ser | Gly | Ser | 20 | 25 | 30 | |
| Pro | Glu | Thr | Gly | Ala | Gly | Ala | Gly | Thr | Val | Ala | Thr | Thr | Pro | Ala | Ser | 35 | 40 | 45 | |
| Ser | Pro | Val | Thr | Leu | Ala | Glu | Thr | Gly | Ser | Thr | Leu | Leu | Tyr | Pro | Leu | 50 | 55 | 60 | |
| Phe | Asn | Leu | Trp | Gly | Pro | Ala | Phe | His | Glu | Arg | Tyr | Pro | Asn | Val | Thr | 65 | 70 | 75 | 80 |
| Ile | Thr | Ala | Gln | Gly | Thr | Gly | Ser | Gly | Ala | Gly | Ile | Ala | Gln | Ala | Ala | 85 | 90 | 95 | |
| Ala | Gly | Thr | Val | Asn | Ile | Gly | Ala | Ser | Asp | Ala | Tyr | Leu | Ser | Glu | Gly | 100 | 105 | 110 | |
| Asp | Met | Ala | Ala | His | Lys | Gly | Leu | Met | Asn | Ile | Ala | Leu | Ala | Ile | Ser | 115 | 120 | 125 | |
| Ala | Gln | Gln | Val | Asn | Tyr | Asn | Leu | Pro | Gly | Val | Ser | Glu | His | Leu | Lys | 130 | 135 | 140 | |
| Leu | Asn | Gly | Lys | Val | Leu | Ala | Ala | Met | Tyr | Gln | Gly | Thr | Ile | Lys | Thr | 145 | 150 | 155 | 160 |
| Trp | Asp | Asp | Pro | Gln | Ile | Ala | Ala | Leu | Asn | Pro | Gly | Val | Asn | Leu | Pro | 165 | 170 | 175 | |
| Gly | Thr | Ala | Val | Val | Pro | Leu | His | Arg | Ser | Asp | Gly | Ser | Gly | Asp | Thr | 180 | 185 | 190 | |
| Phe | Leu | Phe | Thr | Gln | Tyr | Leu | Ser | Lys | Gln | Asp | Pro | Glu | Gly | Trp | Gly | 195 | 200 | 205 | |
| Lys | Ser | Pro | Gly | Phe | Gly | Thr | Thr | Val | Asp | Phe | Pro | Ala | Val | Pro | Gly | 210 | 215 | 220 | |
| Ala | Leu | Gly | Glu | Asn | Gly | Asn | Gly | Gly | Met | Val | Thr | Gly | Cys | Ala | Glu | 225 | 230 | 235 | 240 |

160

```

Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp Gln Ala
      245                      250                      255

Ser Gln Arg Gly Leu Gly Glu Ala Gln Leu Gly Asn Ser Ser Gly Asn
      260                      265                      270

Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala Gly Phe
      275                      280                      285

Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro
      290                      295                      300

Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Gln Tyr Ala Ile Val Asn
      305                      310                      315                      320

Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu
      325                      330                      335

His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp Gln Val
      340                      345                      350

His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu
      355                      360                      365

Ile Ala Thr Ile Ser Ser
      370

```

(2) INFORMATION FOR SEQ ID NO:156:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1777 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

```

GGTCTTGACC ACCACCTGGG TGTGGAAGTC GGTGCCCGCA TTGAAGTCCA GGTACTCGTG      60
GGTGGGGCGG GCGABACAAI AGCGACAAGC ATGCGAGCAG CCGCGGTAGC CTTTGACGGT      120
GTAGCGAAAC GGCAACGCGG CCSCGTTGGG CACCTTGTTC AGCGCTGATT TGCACAACAC      180
CTCGTGGGAG GTGATGCCGT CGAATTGTGG CGCGGGAAGG CTGCGGACCA GGCGGATCCG      240
CTGCAACCCG GCAGCGCCCG TCGTCAACGG GCATCCCGTT CACGCGACAG GCTTGCCCGG      300
CCCACCCCAT ACCATTATTC GAACCAACCT TCTATACTTT GTCAACGCTG GCCGCTACCG      360
AGCGCCUCAC AGGATGTGAT ATGCCATCTC TGCCGCGACA CACAGGAGCC AGGCCATTATC      420
ACAGCAATCG CGCTCGAGCC CTACGGGCGG CCGAATACC TAGAATCCG CCGGAAGCGC      480
ATGCGGTATA TCGACGAAGG CAGGGGTGAC GCCATCGTCT TTCAGCAGCG CAACCCGACG      540
TCTCTTACT TGTGGCGCAA CATCATGCCG CACTTGGGAG GGTGGGGCGG GCTGGTGGCC      600

```

161

```

TCCGATCTGA TCGGATGAGG CCGCTCGGAC AAGCTCAGCC CATCGGGACC CGACCGCTAT      660
AGCTATGGCG AGCAACGAGA CTTTTTGTTC GCGCTCTGGG ATGCGCTCGA CUTCGGGAGC      720
CACGTGGTAC TGGTGTGCA CCACTGGGGC TCGGCGCTCG GCTTCGACTG GCTTAACGAG      780
CATCGCGACC GAGTGCAGGG GAICGCTTC ATGGAAGCGA TCCTCACCCC GATGAGCTGG      840
GCGGACTGGC CGCGCGCCCT CCGGGGTGTC TTCCAGGGTT TCCGATCGCC TCAAGGCGAG      900
CCAATGGCGT TGGAGCAGAA CATCTTTGTC GAACGCGTGC TCGCGGGGCG GATCCTGCGA      960
CAGCTCAGCG ACCAGGAAAT GAACCACTAT CCGCGGCCAT TCGTGAACGG CCGCGAGGAC     1020
CGTCGCCCCA CGTTGTCTG GGCACGAABC CTTCGATCG ACCTGAGGCT CCGCGAGCTC     1080
GTGCGCTTGG TCACCGAGTA CCGGAGCTGG CTCGAGGAAA CCGACATGCC GAAACTGTTG     1140
ATCAACGCGG AGCGCGCGCG GATCATCAGC GCGCGCATCC GTGACTATGT CAGGAGCTGG     1200
CCCAACCGCA CGGAATCAC AGTGGCCGGC GTGCATTTCC TTCAGGAGGA CAGCGATCGC     1260
GTGATATCGT GGGCGGGGCG TCGGCAGGCT CCGCGACCTG GGAGCGCTCT CATTTACGGA     1320
GACCAAGAAAT GTGATTTCCG GCGAAGGCGG CCGCTCTGCT GTCAACTCAT AAGACTTCCT     1380
GCTCGGGGCA GAGATTTCTA GGGAAAAGGG CACCAATCGC AGCGGCTTCC TTGGAACGGA     1440
GGTCGACAAA TATACGTGGC AGGACAAAGG TCTTCCTATT TCGCCAGCGA APTAGTCGCT     1500
GCTTTTCTAT GGGCTCAGTT CGAGGAAGCC GAGCGGATCA CCGGTATCCG APTGAGCTTA     1560
TGGAACGGGT ATCATGAAAG CTTCGAATCA TTGGAACAGC GGGGGCTCCT GCGCCGTCGG     1620
ATCATCCGAC AGGCTCTCTC TCACAACGGC CACATGTACT ACCTGTTACT AGCGCCGAGC     1680
CGCGATCGGG AGGAGGTGCT GGCGCTCTG ACCAGCGAAG GTATAGGCGC GGTCTTTCTAT     1740
TACGTGCCGC TTCAGGATTC GCGGCGCGGG CGTCGCT
                                                                                   1777

```

(2) INFORMATION FOR SEQ ID NO:157:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 324 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(*) SEQUENCE DESCRIPTION: SEQ ID NO:157:

```

GAGATGGAAT CGTACCGGTC TCCTTAGCGG CTCGCTCCCG TGAATGCCCA TATCAGGCAC      60
GGCATGTTT TCGCTGTGGA CTTTCGCCCC ATGCCCGGAC GTTGGTAAAC CCAGGCTTGG     120
ATCAGTAATT CCGGGGACG GTTCCGGGAA GCGGGCCAGG ATGTGCTTGA GCGCGGGGCG     180
CGCGGTGCGG CAGGCGAGCG CTGGATGCTC AGCGCGGCTG CCGCGACGTA GCGAGCGTTT     240

```

GGCGCGTGTG GTCCACAGTG GTACTCCGGT GACGACGGCG CCGGGTGGCT GGCTGAAGAC 300
CGTGACCGAC GCGCGCGATT CAGA 324

(2) INFORMATION FOR SEQ ID NO:158:

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1338 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(*1) SEQUENCE DESCRIPTION: SEQ ID NO:158:

GGGTATACCG CCGTTTCCSC TGGCAGCGGA CCTTACGAC CTGAACCACT TCGCCTGGCG 60
AAGGATTGAC GAACCGCTCG TCGCGCGCT GTGGGTCTC AAGGTGTGG GTGATGTGGT 120
CGATGACCGG CCGCGCACCC GCGCACTACG CGTGAGAGAC GTCTTGCGG CCGCGAGCGA 180
GCAGGACTTC CAGCCCGACT CGATCGGCT GCTGACCCCT CCTGTGCTA TGGCTGCTG 240
GGAGGCTCG GTTCGGAAGC GATTTCGCTT CCTCACTGAC CTGACCGCG ACCGAGAGCG 300
GTGGCGCGCC TGGACGAAAC GGCACCGCG CGAAGTGGAG AACCGGCTGG CGGTCTCTCG 360
GTCTGATCA ACCTGCGGC GATCGTGGG TTCCGCTGGC ACGGTGCGG CTGACGCGG 420
CTGAATCGAC TAGATGAGAG CAGTTGGGCA CGAATCGGC TGTGCTGGTG AGCAAGACAC 480
GAGTACTGTC ATCACTATTG GATGCACTGG ATGACCGGCC TGATTCAGCA GGACCAATGG 540
AACTGCCCCG GCGAAAACGT CTCGAGATG ATCGCGCTCC CCTCGGACC CTGCGGTGCT 600
GGGTTCATTC GGACATCGGT CCGGCTCGCG GGATCGTGGT GAGCGCAGCG CTGAAGGAST 660
GGAGCGCGCG GTTGCACCGG CTGCTGGAGC GCGGCGACAC GGTGCTGCTG CGTAAGGCTG 720
GGATCGGCGA GAACCGCTTC GAGGTGGCG CCGCAGATT CTTCGTTGTC CCGACGCTCG 780
CGCACAGCCA CGCGAGCGCG GTTGGCGCGG AGCACCGCGA CCGCTGCGG CCGGCGGCGG 840
CGACAGCAC CGACGAGTGT GTCTACTCG GCGCGCGAGC GAAAGTTGTT GCGGCACTGC 900
CGTTAACCG GCGAGAGGCT CTGGACGCGA TCGAGGATCT GCACATCTGG ACCGCGGAGT 960
CGGTGCGCGC CGACCGGCTC GACTTTCGSC CCAAGCACAA ACTGGCGCTC TTGTTGTTGT 1020
CGGCGATCCC GCTGCGCGAG CCGGTGCGGC TGGCGCGTAG GCGCGAGTAC GCGGTTTCCA 1080
CGAGCTGGGT GCAGCTGCGG GTGACGCGCA CCGTGGCGGC GCGGCTGAC GACGAGCGCG 1140
CGCTGCGCGA GGTGCGCGC CGGTTCGCG AGCGCGTGGG TTGACTGGGC GGCATCGCTT 1200
GGGTCTGAGC TGTACGCGCA GTGCGCGCTG CGACTGATCT GCTGTGCTTT CGGTCCCTGC 1260
TGGGTCAAT TGACGGCGCG GCGAACAGCA GCATTGCGCG CCGCATCTTC CCGCGCGCGG 1320

GGGCCCCACCG CTACAACC

1338

(2) INFORMATION FOR SEQ ID NO:159:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 321 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:159:

```

CCGGGGGCAC CGSCGGCACC GCGGGTACCG SCGGCAACCG CGCTGACGCG GCTGCTGTGG      60
TGGGCTTCGG CCGCAACGCG GACCGTGGCT TGCTGGCGCG CAAAGGCGGT AACGGCGGAA      120
TACCTGGGCG CCGCGTACCA GCGGGGGTGG CCGGCGACCG CGGCACCGCG GCGAAGGTG      180
GCACCGGCGG TCGCGGCGCG SCGGCAACCG AGCGCGCAC CACCGGCAAT CCGGCGGTA      240
AGGCGGGCGA CCGCGGGATC GCGGTGCGG CGGGGGCGCG CCGCGCGCGG GGCACCGGCA      300
ACGCGGGGCA TCGCGGCAAC C                                     321

```

(2) INFORMATION FOR SEQ ID NO:160:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 492 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:160:

```

GAAGACCCCG CCCCCCATA TCGATCGGCT CGCCGACTAC TTTCGGCGAA CGTGCACGCG      60
GGGCGGTGCG GCTGATCATC ACCGGTGGCT ACSCGCCCAA CGGCACCGGA TGGCTGCTGC      120
CGTTCGCTC CGAACTGCTC ACTTCGGCGC AAGTCCGACG GCACCGCGGA ATCACCAGGG      180
CGGTCCACGA TTCGGGTGCA AAGATCGTGC TGCAAATCCT GCAGGCGGGA CGCTACGCCT      240
ACCACCCACT TCGGCTCAGC GCGTCGCGCA TCAGGCGGCG GATCACCCCG TTTCGTGCGC      300
GAGCACTATC GGTGCGCGCG GTCGAAGCGA CGATCGCGGA TTTCGCGCGC TCGCGCGCAT      360
TGGCCCCGGA TCGCGGTAC GACGGCGTGG AATCATGGG CAGCGAAGGG TATCTGCTCA      420
ATCAGTTCTT GCGCGCGCGC ACCAACAGG GCACCGACTC GTGGGCGCGG ACACCGGCGA      480
ACCGTGGCGG GT                                     492

```

(2) INFORMATION FOR SEQ ID NO:161:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 336 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: Linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:161:

```

Phe Ala Gln His Leu Val Glu Gly Asp Ala Val Glu Leu Trp Arg Ala
1           5           10           15
Asn Ala Ala Asp Gln Ala Asp Pro Leu Gln Pro Gly Ser Ala Arg Arg
20           25           30
Gln Arg Ala Ser Arg Ser Pro Arg Arg Leu Ala Gly Pro Asn Ala Tyr
35           40           45
His Tyr Ser Asn Asn Arg Ser Ile Leu Cys Gln Arg Trp Pro Leu Pro
50           55           60
Ser Ala Ala Gln Asp Val Ile Cys His Leu Cys Pro His Arg Gln Glu
65           70           75           80
Pro Gly Leu Met Thr Ala Phe Gly Val Glu Pro Tyr Gly Gln Pro Lys
85           90           95
Tyr Leu Glu Ile Ala Gly Lys Arg Met Ala Tyr Ile Asp Glu Gly Lys
100          105          110
Gly Asp Ala Ile Val Phe Gln His Gly Asn Pro Thr Ser Ser Tyr Leu
115          120          125
Trp Arg Asn Ile Met Pro His Leu Glu Gly Leu Gly Arg Leu Val Ala
130          135          140
Cys Asp Leu Ile Gly Met Gly Ala Ser Asp Lys Leu Ser Pro Ser Gly
145          150          155          160
Pro Asp Arg Tyr Ser Tyr Gly Glu Gln Arg Asp Phe Leu Phe Ala Leu
165          170          175
Trp Asp Ala Leu Asp Leu Gly Asp His Val Val Leu Val Leu His Asp
180          185          190
Trp Gly Ser Ala Leu Gly Phe Asp Trp Ala Asn Gln His Arg Asp Arg
195          200          205
Val Gln Gly Ile Ala Phe Met Glu Ala Ile Val Thr Pro Met Thr Trp
210          215          220
Ala Asp Trp Pro Pro Ala Val Arg Gly Val Phe Gln Gly Phe Arg Ser
225          230          235          240
Pro Gln Gly Glu Pro Met Ala Leu Glu His Asn Ile Phe Val Glu Arg
245          250          255
Val Leu Pro Gly Ala Ile Leu Arg Gln Leu Ser Asp Glu Gln Met Asn
260          265          270

```

165

```

His Tyr Arg Arg Pro Phe Val Asn Gly Gly Glu Asp Arg Arg Pro Thr
    275                      280                      285

Leu Ser Trp Pro Arg Asn Leu Pro Ile Asp Gly Glu Pro Ala Glu Val
    290                      295                      300

Val Ala Leu Val Asn Glu Tyr Arg Ser Trp Leu Glu Glu Thr Asp Met
    305                      310                      315                      320

Pro Lys Leu Phe Ile Asn Ala Glu Pro Gly Ala Ile Ile Thr Gly Arg
    325                      330                      335

Ile Arg Asp Tyr Val Arg Ser Trp Pro Asn Gln Thr Glu Ile Thr Val
    340                      345                      350

Pro Gly Val His Phe Val Gln Glu Asp Ser Asp Gly Val Val Ser Trp
    355                      360                      365

Ala Gly Ala Arg Gln His Arg Arg Pro Gly Ser Ala Leu Ile Ser Arg
    370                      375                      380

Asp Gln Gln Cys Asp Phe Arg Arg Arg Arg Arg Pro Ala Cys Glu Leu
    385                      390                      395                      400

Ile Arg Leu Pro Ala Pro Gly Arg Asp Ser Gln Gly Lys Gly His Glu
    405                      410                      415

Ser Gln Pro Leu Pro Ser Gln Arg Gly Arg Gln Ile Tyr Val Ala Gly
    420                      425                      430

Gln Arg Ser Ser Tyr Leu Pro Ser Gln Leu Val Ala Ala Phe Leu Trp
    435                      440                      445

Ala Gln Phe Glu Gln Ala Glu Arg Ile Thr Arg Ile Arg Leu Asp Leu
    450                      455                      460

Trp Asn Arg Tyr His Glu Ser Phe Gln Ser Leu Glu Gln Arg Gly Leu
    465                      470                      475                      480

Leu Arg Arg Pro Ile Ile Pro Gln Gly Cys Ser His Asn Ala His Met
    485                      490                      495

Tyr Tyr Val Leu Leu Ala Pro Ser Ala Asp Arg Glu Glu Val Leu Ala
    500                      505                      510

Arg Leu Thr Ser Glu Gly Ile Gly Ala Val Phe His Tyr Val Pro Leu
    515                      520                      525

His Asp Ser Pro Ala Gly Arg Arg
    530                      535

```

(2) INFORMATION FOR SEQ ID NO:162:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 284 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

166

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Glu | Ser | Ala | Pro | Arg | Ser | Pro | Met | Leu | Pro | Ser | Ala | Arg | Pro | Arg | 1 | 5 | 10 | 15 |
| Tyr | Asp | Ala | Ile | Ala | Val | Leu | Leu | Asn | Glu | Met | His | Ala | Gly | His | Cys | 20 | 25 | 30 | |
| Asp | Phe | Gly | Leu | Val | Gly | Pro | Ala | Pro | Asp | Ile | Val | Thr | Asp | Ala | Ala | 35 | 40 | 45 | |
| Gly | Asp | Asp | Arg | Ala | Gly | Leu | Gly | Val | Asp | Glu | Gln | Phe | Arg | His | Val | 50 | 55 | 60 | |
| Gly | Phe | Leu | Glu | Pro | Ala | Pro | Val | Leu | Val | Asp | Gln | Arg | Asp | Asp | Leu | 65 | 70 | 75 | 80 |
| Gly | Gly | Leu | Thr | Val | Asp | Trp | Lys | Val | Ser | Trp | Pro | Arg | Gln | Arg | Gly | 85 | 90 | 95 | |
| Ala | Thr | Val | Leu | Ala | Ala | Val | His | Glu | Trp | Pro | Pro | Ile | Val | Val | His | 100 | 105 | 110 | |
| Phe | Leu | Val | Ala | Glu | Leu | Ser | Gln | Asp | Arg | Pro | Gly | Gln | His | Pro | Phe | 115 | 120 | 125 | |
| Asp | Lys | Asp | Val | Val | Leu | Gln | Arg | His | Trp | Leu | Ala | Leu | Arg | Arg | Ser | 130 | 135 | 140 | |
| Glu | Thr | Leu | Glu | His | Thr | Pro | His | Gly | Arg | Arg | Pro | Val | Arg | Pro | Arg | 145 | 150 | 155 | 160 |
| His | Arg | Gly | Asp | Asp | Arg | Phe | His | Glu | Arg | Asp | Pro | Leu | His | Ser | Val | 165 | 170 | 175 | |
| Ala | Met | Leu | Val | Ser | Pro | Val | Glu | Ala | Glu | Arg | Arg | Ala | Pro | Val | Val | 180 | 185 | 190 | |
| Gln | His | Gln | Tyr | His | Val | Val | Ala | Glu | Val | Glu | Arg | Ile | Pro | Glu | Arg | 195 | 200 | 205 | |
| Glu | Gln | Lys | Val | Ser | Leu | Leu | Ala | Ile | Ala | Ile | Ala | Val | Gly | Ser | Arg | 210 | 215 | 220 | |
| Trp | Ala | Glu | Leu | Val | Arg | Arg | Ala | His | Pro | Asp | Gln | Ile | Ala | Gly | His | 225 | 230 | 235 | 240 |
| Gln | Pro | Ala | Gln | Pro | Phe | Gln | Val | Arg | His | Asp | Val | Ala | Pro | Gln | Val | 245 | 250 | 255 | |
| Arg | Arg | Arg | Gly | Val | Ala | Val | Leu | Lys | Asp | Asp | Gly | Val | Thr | Leu | Ala | 260 | 265 | 270 | |
| Phe | Val | Asp | Ile | Arg | His | Ala | Leu | Pro | Gly | Asp | Phe | 275 | 280 | | | | | | |

(2) INFORMATION FOR SEQ ID NO:163:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 264 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

| | |
|---|-----|
| ATGAACATGT CGTCGGTGGT GGCTCGCAAG GCGTTTGGCG GATTCGCGCG CTACTCCTCC | 60 |
| GCCATGCACG CGATCGCGCG TTCTCCGAT GCGTTGCGCC AAGAGCTCGG GGGTAGCGGA | 120 |
| ATCGCCGTCT CGGTGATCCA CCGGGCGCTG ACCGAGACAC CGCTGTGGC CAACGTGCAC | 180 |
| CCCGCGGACA TCGCGCGCGG GTTTCGCGAC CTCACGCGCA TTCCCGTCCA CTGGGTGCGG | 240 |
| GCAGCGGTGC TTGACGGTGT GGCG | 264 |

(2) INFORMATION FOR SEQ ID NO:164:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1171 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

| | |
|---|-----|
| TACTCGCGCA CGATGACGTC GCGTCCAGG CCGACCGCTT CAAGCACCAG CCGGACCACG | 60 |
| AAGCCCGGTGC GATCCTTACC CGCGAAGCAG TGGGTGAGCA CCGGGCGTCC GCGGCAAGC | 120 |
| AGTGTGACCA CACGATGTAG CCGCGGTGT GCTCCATTGC CGCTTGGAA TTGGGATAC | 180 |
| TCCTCGGTCA TGTAGCGGCT GCGCGGTCA TTTATCGACT GCGTGGATT CCGGGACTCG | 240 |
| CCGTTGGAAC CGTCAATTGT TACGAGCCTC TTGAATGCG TTTCGTGCG CGCTGAGTCG | 300 |
| TCGGCGTCAT CATCGGCGAG GTCGGGGAC GGCAGCAGGT GCACCTCAT GCGTCCGGA | 360 |
| ACCGTCCCTG GACCGCGCG GCGAAGCTCC CCGGACGACC GCAGTCCGC AACGTCCGTG | 420 |
| ATCCCCAGGC GCGCGAGCGT TCGCCCTCGT GCGGAATTGG GCACGAGGCT GCGGAGCCAC | 480 |
| CGGGCATTAC CAAGCAAGCC TTGCCCAGTA CGGATCGTCA CTTCGCGATC CGGCGACCA | 540 |
| ATCTCCTGCG CGCCATCCT CAGATCCGCG TCGTGGGTTG ACAGGACCG CCGCAGATGT | 600 |
| GCGAGCGGCT ATCGGAGATT GAGCCGCGCA CGCATTCCTT CAATCGCTGC GCGCTGCGCG | 660 |
| ACTATTGCA CTTTCGCGG GTGCGGTAT TCAGCAGCA TCGGAGTCTC GACGAACTCG | 720 |
| CCGACGTAA CCGCGGGCT AGCTCCCGCG GTGAGCGGGA GGTTCGCGG GTGATCTTTC | 780 |
| CGGCGAGCT CAGAGCGGT GATCCAGCG TTGCGCGTGC CGGCGGGAG GCGGATCAGC | 840 |

168

```

TTATCGACCT CGGGGTATGC CGACGGCAAG CTGGGGGGT TCCTCGAGGT CAGAACTCC      900
ACCATCGGCA CCGGCACCAA GGTGCGGCAC CTGACCTAGC TCGGCGACGC CGACATCGGC      960
GAGTACAGCA ACATCGGGCG CTCCAGCGTG TTCCTCAACT ACGACGGTAC GTCCAAACGG      1020
CGCACCACCG TCGGTTGCGA CGTACGGACC GGGTCCGACA CCGTGTTCGT GGGCCCACTA      1080
ACCATCGGCG ACGGCGCGTA TACCGGGGCG GGCACAGTGG TCGGCGAGGA TGTCCCGCGC      1140
GGGGCGCTGG CAGTGTCCGC GGGTCCGCAA C

```

(2) INFORMATION FOR SEQ ID NO:165:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 227 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:165:

```

GCAAGGCGCG CACCGGCGCG GCGGGCATGA ACAGCCTCGA CCGGCTGCTA GCGGCGCAAG      60
ACGGGCGGCA AGGCGGCACC GCGGGCACCG GCGGCAACGC CGGCGCGGCG GGCACCTAGT      120
TCACCCAAAG CGCGGACGCG AACGCGGCGA ACGGCGGTGA CGGCGGGGTC GCGGGCAACG      180
GCGGAACCGG CGGAACGGC GCAGACAACA CCGCCACCGC CGCGGCG

```

(2) INFORMATION FOR SEQ ID NO:166:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 304 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:166:

```

CCTCGGCAAC ATGGGCGGCG AGGGCGGTAG CGGTGCGGCG GCTCTACCC CAGGCGCCAA      60
GGGCGGCGAC GGCTTCACTC CAGCCAGCGG CGGCGACGCG GCGGACGGCG GCACCGGCGG      120
CAACTCCCAA CTGGTCGGCG GCAACGGCGG CGACGGCGCG AATGGCGGCA ACGGCGGCGG      180
CGCGGCGACG GCGGCGCAAC GCGGCGGCGG CGGCGACGCG GCTTTGGTG GCATGAGTGC      240
CAACGGCACC AACCTGCTG AAAACGGGCG AAACGTTAAC CCGGCGGCGA ACCTGCGGCG      300
CGGC

```

(2) INFORMATION FOR SEQ ID NO:167:

(1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1439 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:167:

```

GTGGGACGCT GCCGAGGCTC TATAACAAGG ACAACATCGA CCAGCGCCCG CTCGGTGAGC      60
TCATCGACCT ATTTAACAGT GCGCGCTTCA GCGGCGAGGG CCAGCACCCT GCGCGGATC      120
TGATGGGTGA GCTCTACGAA TACTTCCTCG GCAATTTGGC TCGCGCGGAA GGGGAGCGGG      180
GTGGCGAGTT CTTTACCCCG CCCAGCGTGG TCAAGGTGAT CCGGAGGTG CTGGAGCCGT      240
CGAGTGGGCG GGTGTATGAC CCGTCTGCG GTTCGGGAGG CATCTTTGTG CAGACCGAGA      300
AGTTCACTCA CGAACACGAC GCGGATCGGA AGGATGTCTC GATCTATGGC CAGGAAAGCA      360
TTGAGGAGAC CTGGCGGATG GCGAAGATGA ACCTCGCCAT CCACCGCATC GACACACAGG      420
GGCTCGGCGC CGGATCGAGT GATACCTTCC CCCCGGACCA GCACCCGGAC GTGCAGATGG      480
ACTACGTCAT GCGCANTCCG CCGTTCARCA TCAAAGACTG GCGCGGCAAC GAGGAAGACC      540
CAGGCTGGCG CTTCGGTGTT CCGCGCGCCA ATAAGGCCAA CTACGGATGG ATTCAGCACA      600
TGCTGTACAA CTTGGCGCCG GGAGGTCCGG CGGGCGTGGT GATGGCCAAC GGGTCGATGT      660
CGTCGAACCT CAACGGCAAG GGGGATATTC GCGCGCAAAAT CCGGAGGCGG GATTTGCTTT      720
CCTGCATGGT CGCGTTACCC ACCCAGCTGT TCGGAGGAC CGGAATCCCG GTGTGCTGT      780
GCTTTTTCGC CAAAACAAG GCGGCAGGTA AGCAAGGCTC TATCAACCGG TCGCGGCGAG      840
TGCTGTTTAT CGACCTCGT GAACTGGGCG ACCTAGTGGG CCGGGCCGAG CCGGCGCTGA      900
CCAGCGAGGA GATCGTCCCG ATCGGGGATA CTTTCCACGC GAGCAGGACC ACCCGCAACG      960
CCGCGTCCGG TGSTGCCGCG GGTAAATGGG GCACTGGCCT CAGCGGCGCG GCGGCTGCTG      1020
GCGGGGCGCG CGGCAACCGG GGTGTCCCGG GCGTGTCTT CCGCAACGCT GTGGGCGGCG      1080
ACGGCGGCAA CGCGGCAAC GCGGTCACG GCGGCGACGG CAGGAGGCGC GCGGCGGCG      1140
GCAAGGCGCG CAGCGCAGC AGCGGTGGCG CCGCGGCTC AGGCGTGGTC AACGTCACCG      1200
CCGCGCACCG CGGCAACCGC GCGAATGGCG CCAACGGCG CAGCGCTCC GCGGGCGCG      1260
GCGGCGAGGG CGGTGCGGG GGCAGCGCG GCAACGGCG CAGCGGCGC GTGCGCACCG      1320
GCGGCGGCG CCGCAAGGCG GCGAAGGCG CAGCGGCTC CCGCAGCGC TCAGGCGTCA      1380
TCAACGTCAC CCGCGGCGC GCGGCGAGG GCGGCAATCG CCGCAACCGC GCGAACCG      1439

```

(2) INFORMATION FOR SEQ ID NO:168:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 329 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

```

GGGCGCGCGG GCGCGGATT TCTCGTGCTT TGATGTGCGG TGGGGATAAC GCGCGTGATG      60
GTGGTAACCG CGGGATGGGC GGGGCTGGCG GGGCTGGCGG CCGCGGCGCG GCGCGCGCGC      120
TGATCAGCCT GCTGGGCGGC CAAGGCGCGC GCGGCGCGCG CGGGACCGGC GGGGCGCGCG      180
GTGTTGGCGG TGACGCGCGG GCGGCGCGGC CCGGCAACCA GGCCTTCACG GCGGGTGGCG      240
GCGGCGCGCG CGGCTGATC AGCCTGCTGG GCGGCCAAGG CCGCGGCGCG GCGGCGCGGA      300
CGGCGCGGCG CCGCGGTGTT GCGGTCAC                                     329

```

(2) INFORMATION FOR SEQ ID NO:169:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 80 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

```

GCAACGGTGG CAACGCGCGC ACCGCAACA CCGTGGGAT GCGCGGCGGT AACTGTGCTG      60
CGGCGGCGGT GATCGGCAAC                                     80

```

(2) INFORMATION FOR SEQ ID NO:170:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 392 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

```

GGGCTGTGTC GCACTCACAC GCGCGCATTC GCGACGTTG GCGGCGCAAT ATCCAGCTCA      60
AGGCGTACTA CTTACCGTGG GAGGACCGCC GCATCAAGGT GCGGGTCAGC GCGCAAGGAA      120
TCAAGGTCAAT CGACCGGAG GGCATCGAG GCGTCTCTGG CCGGCTGCG GCGGATCCG      180

```

171

```

CCCCGUCUCA CTTCCGCGGC CAAGCGGGCT CATCGCTCCG AACGGCGGGG ATCCTGTGAG      240
CACAACTGAT GCGCGGCAAC GAGATTCGTC CAATTTGCAA GCGGTGTTCC ACCCCAGGGA      300
CCGGTTATAC GATGTGCAAC CTATGTCAC TCGAAGAAC GGCATAACGA TCCCGTGATC      360
CGCCGACAGC CCACCACTGC AAGACCGTTA CA                               392

```

(2) INFORMATION FOR SEQ ID NO:171:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 535 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:171:

```

ACCGGCGCCA CCGGCGGCAC CGGGTTGCCC GGTGGGCGCG GCGGGGCTGG CCGGCAGGGC      60
GGTATCAGCG GTGCGCGGCG CACCAACGGC TCTGGTGGCG CTGGCGGCAC CGGCGGACAA      120
GGCGGCGCGG GGGGCGCTGG CGGGGCGGGC GCGGATAACC CCACCGGCAT CCGCGGCGCC      180
GGCGGCACCG GCGGCACCGG CGGAGCGGGC GGAGCGGGCG GGGCGGCTGG CGCATCTGGT      240
ACCGGCGGCA CCGGCGGCAC GGTGGGCGCG GTGGGTAAAG CCGGGATCGG CGGTACGGGC      300
GGTACGGGTG GTGTGGGTGG TGCTGGTGGT GCAGGTGCGG CTGCGGCGCG TGCGAGCAGC      360
GCTACCGGTG GCGCGGGATT CCGCGGCGGC GCGGCGCGAG AAGCGCGACC GCGCGGCAAC      420
AGCGGTGTGG GCGGCACCAA CGGCTCGGGC GCGCGCGGCG GTGCGAGCGG CAGCGGCGGC      480
ACCGGAGGTG CCGGCGGGTC CCGGCGGGAC AACCCACCG GTGCTGGTTT CCGCG      535

```

(2) INFORMATION FOR SEQ ID NO:172:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 690 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:172:

```

CCGACGTGCG CCGGCGGATA CCGGGGTCAC CGACTACTAC ATCATCCGCA CCGAGAATCG      60
GCGGCTGCTG CAACCGCTGC GGGCGGTGCG GGTATTCGGA GATCCGCTGG CCGACCTGAT      120
CCAGCCGAAC CTGAAGGTGA TGCTCAACCT GGGCTACGGC GACCCGAAGT ACGGCTACTC      180
GACGAGCTAC GCGGATGTGC GAACGCGGTT CCGGCTGTGG CCGAACGTGC CGGCTCAGGT      240

```

172

```

CATCCCGGAT GCGCTGGCCG CCGGAACACA AGAAGGCATC CTTGACTTCA CCGCCGACCT      300
GCAGGCGCTG TCGGCGCAC CCGTCACCGT CCGGCAGATC CAGCTGCTCC AACCCGCCSA      360
TCTGGTGGCC GCGGTGGCCG CCGCACCGAC GCGGGCCGAG GTGGTGAACA CGCTCGCCAG      420
GATCATCTCA ACCAACTACG CCGTCTCTGT GCGCACCGTG GACATCGCCG TCCTCTGGTC      480
ACCACCTTGC CGCTGTACAC CAGCCAACTG TTGTCAGGC AACTCGCTGC GGGCAATCTG      540
ATCAACGCGA TCGGCIATCG CCTGGCGGCC ACGGTAGGTT TAGGCACGAT CGATAGCGGG      600
CGGCTGGAAT TTGCTCACCC TCCTCGCGGC GCGCTCGGAC ACGGTTCGAA ACATCGAGGG      660
CCTCGTCACC TAACGGATTC CCGACGGCAT      690

```

(2) INFORMATION FOR SEQ ID NO:173:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 407 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(*1) SEQUENCE DESCRIPTION: SEQ ID NO:173:

```

ACGGTGACGG CGGTACTGGC GCGGCCCAAG GCGGCACCG CCGGAATCCC GGSTGGCTCT      60
TGGGCACAGC CCGGGGTGGC GCGAACGGTG GCGCCGCGAG CACCGGTACT GCAGGTGGCG      120
GCTCTGGGGG CACCGGCGGC GACGGCGGGA CCGGCGCGCG TGGCGGCTG TTAATGGCGG      180
CCGCGCGCGG CCGGCACGGT GGCACCTGGC GCGCGGCGCG TCGCGGTGTC GACGGTGGCG      240
GCGCCGCGCG GCGCGGCGCG GCGCGCGGCA ACGGCGCGCG CCGGGGTCAA GCGCGCTGTC      300
TGTTCGGGGC CCGCGGCGAC GCGGAGCGCG GCGGCTACCG CCGCGATGCG GGTGGCGGCG      360
GTGACGGCTT CGACGGCAGG ATGCGCGGCG TGGGTGGTAC CCGTGGC      407

```

(2) INFORMATION FOR SEQ ID NO:174:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 466 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(*1) SEQUENCE DESCRIPTION: SEQ ID NO:174:

```

GATCGGTGAG CGCATGCCCC TCGGCGGCAA GCGATTCCGC GGTCTCACCG AAGAACATCG      60
TGCACGCGCG GCGCGGGACC AGCCCGCTGC GCTGCGGCGC GTCGAACGCC TCCAGCAGGC      120

```

173

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| ACAGCCAGTC | CTTGGCGGCC | TGGAGGCGCA | ACACGTGGGT | GTCACCGGTG | TAGATCGCCG | 180 |
| GGATGCGCGC | CTCGGCGCAC | GCATTCGGGC | ACGCCCCGCG | GTCCTTGTGA | TGCTCGACGA | 240 |
| TCACCGCGAT | GTCTGCGGCG | ACCACGGGCG | GCCCGGCGAA | GGTGCCCCCG | CTGGCCACTA | 300 |
| GCGCGCGGAC | GTGGGCGGCG | AGGTGGTGGG | GGATGTGGCG | GCGCAGCGGT | CGGGCGCGAC | 360 |
| GCGCGAAGAA | CGACCCCTCA | CCAGCTGGG | TCCCGCTGGC | ATATCCCTTG | CGTCCCTGGC | 420 |
| CGATATTGGA | CGCGCATGCC | CGACCGCGGT | ACAGGCGCGC | CACCAACG | | 468 |

(2) INFORMATION FOR SEQ ID NO:175:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 219 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:175:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| GGTGTAACG | GCGGCCAGGG | TGGCATCGGC | GCGCGCGCGG | ACAGAGGCGC | CGACGGCGCC | 60 |
| GCGCCCAATG | CTAACGGGCG | AAACGGCGAG | AACGGCGGTA | GCGTGGTAA | CGGTGGCGAC | 120 |
| GCGCGCGCGG | GCGGCAATGG | GCGCGCGGCG | GCGACCGCGC | AGGCGGCGCG | GTACACCGAC | 180 |
| GCGGCCACGG | GCACCGCGCG | CGACGGCGGC | AACGGCGCGC | | | 219 |

(2) INFORMATION FOR SEQ ID NO:176:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 494 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:176:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| TAGCTCCGCG | GAGGGCGGCA | AGGGCGGCGA | CGTGGCGCGC | GGGGGTGACG | GCCTCGGCGC | 60 |
| CAACAGTTCC | GTCACCCAAE | GCGGCAGCGG | CGGTGGCGCG | GCGCGCGGCG | GGCGCGGCGG | 120 |
| CAGCGGCTTY | TTGGCGGCGA | AGGGCGGCTT | CGCGCGCGAC | GCGGTCAGG | GCGGCCCCAA | 180 |
| CGCGCGCGGT | ACGTGCGGCA | CGGTGGCGCG | TGGCGGCGCG | AACGGCGGTA | TGGCGGCGCG | 240 |
| GCGCGGCGAC | GCGGTCTTTG | CGGTGGCGCG | CGGCGAGGCG | GCGCTCGGTA | GCGAGGCGCG | 300 |
| CAATGGCGGG | GCTTCACCG | GCGGCAACCG | CGGCTTTGGC | GCGCGGCGCG | GTGGCGGAGG | 360 |
| CAACGCGCGG | GCTCGTGCGG | AATCGGGGCT | CACCATGGAC | AGCGCGGCGA | AGTTGGCTGC | 420 |

CATCGCATCA GGGCGCTACT GCGCCGAACA CTTGGAACAT CACCGGAGTT AGCGGGGGCGC 480
 ATTTCCTGAT CACC 494

(2) INFORMATION FOR SEQ ID NO:177:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 228 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:177:

GGGCGGCTGG TGCCGCGGGC CAGCTTTTCA GCGCGGAGG CGCGGCGGCT GCGTTGGGG 60
 TTGGCGGCAC CGCGGGCCAG GGTGGGCTG GCGGTGCCCG AGCGGCGGCG GCGGACGCTT 120
 CCGCCAGCAC AGTCTAAC CCCTGTACCG GGTTCCTGG CCGGGGCGCG GCGGTCCGGG 180
 GCGAGAGCGG CAACGGCATT GCGGGCGGCA TCAACGCTC 220

(2) INFORMATION FOR SEQ ID NO:178:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 368 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:178:

ATGCGGGCAA CGGGGGCCCC GCGGTGCTG GCGGGGCGCG CGACTACAAT TTCCAACGGC 60
 GGGCAGGCTG GTGCCGCTG CCAAGGCGCG CAAGGCGGCG TGCGGGGCG AAGCACCACC 120
 TGATCGGCTT AGCGGCACCG GCGAAAGCG ATCCAACAG CGACGATGCG GCGTTCTTGG 180
 CCGGTTTGGG CCAGGCGGCG ATCAGCTAG CTGAGCCAG CCAGGCCATA ACGGCGGCGA 240
 AGGCGATGTC TGGCTGTGT GCTAACGCG TACAGGTCT ACAGCTGCTC GCGGACCTGC 300
 GGGACTACAA TCCCGGCTG ACCATGGACA GCGCGGCGAA GTTCGCTGCG ATCGCATCAG 360
 GCGGCTACTG CCGCGAACAC CTGGAACA 398

(2) INFORMATION FOR SEQ ID NO:179:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 400 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:179:

| | |
|--|-----|
| GCAAAGGCGG CACCGGCGGG GCGGSCATGA ACAGGCTCGA CCGGCTGCTA GCGGCCCCAAG | 60 |
| ACGGCGGGCCA AGCGGGCACC GCGGSCACCG GCGGCAACGC CCGGCGCGGC GGC&CCAGCT | 120 |
| TCACCCAAAGG CCGCGACGGC AACGCGGGCA ACAGGCTGA CCGGCGGGTC GCGGSCAACC | 180 |
| GCGGAAGCGG CCGAAGCGGC GCAGACAACA CCACACCGGC CCGGCGCGGC ACCACAGGGC | 240 |
| GCGACGGCGG GCGCGCGGG GCGGCGGAA CCGGCGGAAC CCGGCGAGCC GCGGSCACCG | 300 |
| GCACCGGCGG CCAACAAGGC AACGCGGGCA ACAGGCGCAC CCGGCGCAAA GCGGSCACCG | 360 |
| GCGGCGACCG TSCACTCTCA GCGGSCACCG GTGGTGGCGG | 480 |

(2) INFORMATION FOR SEQ ID NO:180:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 538 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:180:

| | |
|---|-----|
| GCGACGGCGG GCAACGGCGG CATCGCGGGC ATTGGCGGGC AACGGCCTT CCGGACGGGC | 60 |
| AGCGGCAAGG GCGGCGAAG GCGGCGAGCG GCGCAACGGC GCGAAGCGG GCATGGGCGG | 120 |
| CAACAGCGGC ACGGCGAGCG GCGACGGCGG TGCGGCGGGC AACGGCGGCG CCGCGGGCAC | 180 |
| GCGGCGCACC GCGGCGAGG GCGGCGCTAC CGTACTGGC GCGACCGCGG GCAGCGGTGG | 240 |
| CACCGCGCGT CACGGCGCTA ACGGCGGCAA CCGAGCAGAT AACACCGCAA ACATGACTGC | 300 |
| GCAGCGGGGC GGTGACGGG GCAACGGCGG CCGCGGTGGC TTGGCGGGCG GGGCGGGGGC | 360 |
| CGCGGCGCGT GCGTGAAGG CTGGCGGCAA CCGCACGGGC GCGCAAGCGG GCGCGGGCGG | 420 |
| CGATGGCGGC AACGGGGCCA TCGGCGGCCA CCGCGGCTC ACTGACGACC CCGGCGGCAA | 480 |
| CGGGGGCACC GCGGCGAAG GCGGCAACCG CCGCACCGGC GCGGCGGCAA TCGGCGGC | 538 |

(2) INFORMATION FOR SEQ ID NO:181:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 239 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| GGGCGGGTGG | TGCGCGGGGC | CAGCTCTTCA | GCSCCGGAGG | CGCGGCGGGT | GCCGTTGGGG | 60 |
| TTGGCGGCAC | CGGCGGCCAG | GGTGGGGCTG | GCCTGCGCGG | AGCGGCGGGC | GCCGAGCGCC | 120 |
| CGCCAGGCAC | AGGTCTAACC | GCTGGTACCG | GTTCGCTGG | CGGGGCGGGC | GGCGTCGGCG | 180 |
| GCCACGCGCG | CAACGCCAAT | GCGGCGGGCA | TCAACGGCTC | CGTGGTGGC | GGCGGCACC | 239 |

(2) INFORMATION FOR SEQ ID NO:182:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 995 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|-----|
| AGCAGCGCTA | CGGTTGGGCG | CGGGTTGGCG | GGCGGCGCGG | GGGAGAAGG | CGGAGCGGGC | 60 |
| GGCAACAGCG | GTGTGGGCGG | CACCAACGGC | TCGGGCGGGG | CGGCGGTTGC | AGGCGGCAAG | 120 |
| GGGCGCACCG | GAGGTGCCGG | CGGTTCCGCG | GGGACAAAC | CCGCGGTTGC | TGGTTTGGCG | 180 |
| GCTGGCGCGG | GCGGCACAGG | TGGCGCGGCG | GGCGCGCGCG | GGGCGGCGCG | GGCGACCGGT | 240 |
| ACCGCGCGCA | CGGCGGCGGT | TGTCGGCGCG | ACCGGTAGTG | CAGGATCGG | CGGGCGCGCG | 300 |
| GGCGCGCGCG | GTGACGCGCG | CGATGGGCGG | AGCGGTCTCG | GGTGGGCGGT | CTCGGCGCTT | 360 |
| GACGCGCGCG | AAGCGCGCGA | AGGCGGGGCG | GGCGGCGCGG | CGGCGGCGCG | CGGCATCAAC | 420 |
| GGGCGCGCGG | GGGCGGCGCG | CAACGCGCGG | GACGCGCGCG | ACGCGGCGAG | CGGTGCGCGA | 480 |
| GGTCTCGCGG | ACAACGCGCG | GGTGGGCGGT | GAGGTGGGGG | CGGTTGGCGG | CGCGGCGAAG | 540 |
| GGCGGCAAGG | CGGCGGTCGG | CCTGACAGCG | AAGCGCGGCG | ACGCGGCGCG | CGCGGCGAAT | 600 |
| GGGCGCAAGG | GGGCGGCGCG | CGGTGCTGGG | GGGCGGCGCG | ACAACAAATT | CAACGCGCGG | 660 |
| CAGGTTGGTG | CGGCGGCGCA | AGGCGGCGCA | GGGCGGCTTG | GGGCGGCGAG | CACCGCTTGA | 720 |
| TGGGCTTATC | CGGCGGCGCG | AAAGCGGATC | CAACAGGCGA | CGATGCGGCG | TTCCTTGGCG | 780 |
| CGTTGGACCA | GGCGGCGATC | ACCTAGGCTG | ACCGAGGCGA | CGGCTATACG | GGCGGCGAAG | 840 |
| CGATGTTGCG | GCTGTGCTGT | AACCGCGTAA | CAGGTCTACA | GCTGCTGCGG | GAGCTGCGCG | 900 |
| AATACAAATC | CGGCGTGAGC | ATGAGCGAGG | CGGCGCAAGT | CGTGGCGATC | CGATCGGCGG | 960 |
| CGTACTGCGG | CGAACAAGCTG | GAACA | | | | 995 |

(2) INFORMATION FOR SEQ ID NO:183:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2138 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:103:

```

CGGCACGAGG ATUGGTACCC CGCGGCATCG GGAGCTGCCC ATTCCGCCGG TTCCCCACC      60
CGAGGAAAGC CGCTACCAGA TGGCGCTGCC GAAGTAGGGC GATCCGTTCC CGATCCCGGC      120
ATGAACGGGC GGCATCAANT TASTGCAGGA ACCTTTCAGT TTAGCGACGA TAATGGGTAT      180
AGCACTAAGG AGGATGATCC GATATGATCC AGTCGCAGAC CGTGACGGTG GATCAGCAAG      240
AGATTTTGA AAGGGCCAC GAGGTGGAGG CCCCATGCC GGACCCACC ACTGATGTCC      300
CCATCACACC GTCCGACTC ACGGCGECTA AAAACGCCGC CCAACAGCTG GTATTGTCCG      360
CCGACAACAT GCGGAATAC CTGGCGGCCG GTCCCAAGA GCGGACGGT CTGGCGACCT      420
CGCTGCGCAA CGCGGCAAG GCGTATGGCG AGGTTGATGA GGAGGTTCCG ACCGCGCTCG      480
ACAACGACGG CGAAGGAAC GTGCGAGCAG AATCGGCCGG GCGCGTCGGA GCGGACASTT      540
CGGCGGAAC TACCGATACG CCGAGGCTGG CCACGGCGCG TGAACCCAC TTGATGGATC      600
TCAAGAAGC GCGAAGGAAG CTGGAAGCCG GCGACCAAGG CGCATCGCTC GCGCACTTTC      660
CGGATGGGTG GAACACTTTC AACCTGACGC TCGAAGGCGA GGTCAAGCGG TTCCGGGGGT      720
TTGACAAC TGAGGGCGAT GCGGCTACCG CTTGCGAGGC TTGCTCGAT CAACAACGGC      780
AATGGATACT CCACATGGCC AAATTGAGCG CTGCGATGCC CAGGCGGCT CAATATGTCC      840
CGCAGCTGCA CGTGTGGGCT AGGCGGGAAC ATCCGACTTA TGAAGACATA GTGGGCTCCG      900
AACGGCTTTA CGCGGAAAAC CCTTCGGGCC GCGGCCAANT TCTCCCGGTG TACGCGGAST      960
ATCAGCAGAG GTCGGAGAAG GTGCTGACCG AATACAACAA CAAGGCAGCC CTGGAACCGG      1020
TAACCCGCGC GAAGCTTCC CCGGCCATCA AGATCGACCC CCCCCGCTT CCGCAAGAGC      1080
AGGGATTCAT CCTTGGCTTC CTGATGCCCG CGTCTGACCG CTCGGGTGTC ACTCCGGTGA      1140
CGGGATGCC AGCGGCACCG ATGCTTCCCG CTACCGGATC GCGGGGTGGT GCGCTCCCGG      1200
CTEACACGGC GCGCGAGCTG ACCTCGGCTG GCGCGGAAGC CGCAGCGCTC TCGGCGGACG      1260
TGGGGTCA AAGGGCATCG CTCGGTGGCG GTGAGGCGCG CCGGCTGCCG TCGGCGCGGT      1320
TGGGATCCCG GATCGGGGCG GCGGATCGG TCGGCGCGCG TCGGCTGGT GACATTGCCG      1380
GCTTAGGCCA GGAAGGGCG GCGGCGGCG CCGGCTCGG CCGGCTGGC ATGGAATGC      1440
CGATGGGTG CCGCATCAG GGACAAGGGG GCGGCAAGTC CAAGGTTCT CAGCAGGAAG      1500

```

178

```

ACGAGGCGCT CTACACCGAG GATCGGGCAT GGACCGAGGC CGTCATTGGT AACCGTCGGC 1560
GCCAGGACAG TAAGGAGTCC AAGTGAGCAT GGACCAATTG GACCCGCATG TCGCCCGGGC 1620
GTTGACGCTG GCGGCGCGGT TTCAGTCGGC CCTAGACGGG ACCCTCAATC AGATGAACAA 1680
CGGATCCTTC CCGGCCACCG ACGAAGCCGA GACCGTCGAA GTGACGATCA ATGGGCACCA 1740
GTGGCTCACC GGCTTGGCCA TCGAAGATGG TTGCTGAAG AAGCTGGGTG CCGAGCGGGT 1800
GGCTCAGCGG GTCACCGAGG CGCTGCACAA TCCGCAGGCC GCGGCGTCCG CGTATAACGA 1860
CGCGGCGGGC GAGCAGCTGA CCGCTGGCTT ATGGGCCATG TCCCGCGCGA TGAACGAAGG 1920
AATGGCCTAA GCCCATTGTT GCGGTGGTAG CGACTACGCA CCGAATGAGC GCGGCAATGC 1980
GGTCATTGAG CCGCCCCGAC ACGGCGTGAG TACGCTTTST CAATGTTTTG ACATGGATTC 2040
GCGGGGTTTG GAGGCGCCCA TACTCTGGT CGGCAATATT GCCCGAGCTA GTTGGTCTTA 2100
GTTTCGGTFA CGCTGGTTAA TTATGACGTC CATTACCA 2138

```

(2) INFORMATION FOR SEQ ID NO:184:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 460 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

```

Met Thr Gln Ser Gln Thr Val Thr Val Asp Gln Gln Glu Ile Leu Asn
1           5           10           15

Arg Ala Asn Glu Val Glu Ala Pro Met Ala Asp Pro Pro Thr Asp Val
20           25           30

Pro Ile Thr Pro Cys Glu Leu Thr Ala Ala Lys Asn Ala Ala Gln Gln
35           40           45

Leu Val Leu Ser Ala Asp Asn Met Arg Glu Tyr Leu Ala Ala Gly Ala
50           55           60

Lys Glu Arg Glu Arg Leu Ala Thr Ser Leu Arg Asn Ala Ala Lys Ala
65           70           75           80

Tyr Gly Glu Val Asp Glu Glu Ala Ala Thr Ala Leu Asp Asn Asp Gly
85           90           95

Glu Gly Thr Val Gln Ala Glu Ser Ala Gly Ala Val Gly Gly Asp Ser
100          105          110

Ser Ala Glu Leu Thr Asp Thr Pro Arg Val Ala Thr Ala Gly Glu Pro
115          120          125

Asn Phe Met Asp Leu Lys Glu Ala Ala Arg Lys Leu Glu Thr Gly Asp
130          135          140

```

179

Gln Gly Ala Ser Leu Ala His Phe Ala Asp Gly Trp Asn Thr Phe Asn
 145 150 155 160
 Leu Thr Leu Gln Gly Asp Val Lys Arg Phe Arg Gly Phe Asp Asn Trp
 165 170 175
 Glu Gly Asp Ala Ala Thr Ala Cys Glu Ala Ser Leu Asp Gln Gln Arg
 180 185 190
 Gln Trp Ile Leu His Met Ala Lys Leu Ser Ala Ala Met Ala Lys Gln
 195 200 205
 Ala Gln Tyr Val Ala Gln Leu His Val Trp Ala Arg Arg Glu His Pro
 210 215 220
 Thr Tyr Glu Asp Ile Val Gly Leu Glu Arg Leu Tyr Ala Glu Asn Pro
 225 230 235 240
 Ser Ala Arg Asp Gln Ile Leu Pro Val Tyr Ala Glu Tyr Gln Gln Arg
 245 250 255
 Ser Glu Lys Val Leu Thr Glu Tyr Asn Asn Lys Ala Ala Leu Glu Pro
 260 265 270
 Val Asn Pro Pro Lys Pro Pro Pro Ala Ile Lys Ile Asp Pro Pro Pro
 275 280 285
 Pro Pro Glu Glu Glu Gly Leu Ile Pro Gly Phe Leu Met Pro Pro Ser
 290 295 300
 Asp Gly Ser Gly Val Thr Pro Gly Thr Gly Met Pro Ala Ala Pro Met
 305 310 315 320
 Val Pro Pro Thr Gly Ser Pro Gly Gly Gly Leu Pro Ala Asp Thr Ala
 325 330 335
 Ala Gln Leu Thr Ser Ala Gly Arg Glu Ala Ala Ala Leu Ser Gly Asp
 340 345 350
 Val Ala Val Lys Ala Ala Ser Leu Gly Gly Gly Gly Gly Gly Val
 355 360 365
 Pro Ser Ala Pro Leu Gly Ser Ala Ile Gly Gly Ala Glu Ser Val Arg
 370 375 380
 Pro Ala Gly Ala Gly Asp Ile Ala Gly Leu Gly Gln Gly Arg Ala Gly
 385 390 395 400
 Gly Gly Ala Ala Leu Gly Gly Gly Gly Met Gly Met Pro Met Gly Ala
 405 410 415
 Ala His Gln Gly Gln Gly Gly Ala Lys Ser Lys Gly Ser Gln Gln Gln
 420 425 430
 Asp Glu Ala Leu Tyr Thr Glu Asp Arg Ala Trp Thr Glu Ala Val Ile
 435 440 445
 Gly Asn Arg Arg Arg Gln Asp Ser Lys Glu Ser Lys
 450 455 460

(2) INFORMATION FOR SEQ ID NO:185:

180

(1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 277 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Gly | Asn | Val | Thr | Ser | Ala | Ser | Gly | Pro | His | Arg | Phe | Gly | Ala | Pro | 1 | 5 | 10 | 15 |
| Asp | Arg | Gly | Ser | Gln | Arg | Arg | Arg | Arg | His | Pro | Ala | Ala | Ser | Thr | Ala | 20 | 25 | 30 | |
| Thr | Glu | Arg | Cys | Arg | Phe | Asp | Arg | His | Val | Ala | Arg | Gln | Arg | Cys | Gly | 35 | 40 | 45 | |
| Phe | Pro | Pro | Ser | Arg | Arg | Gln | Leu | Arg | Arg | Arg | Val | Ser | Arg | Glu | Ala | 50 | 55 | 60 | |
| Thr | Thr | Arg | Arg | Ser | Gly | Arg | Arg | Asn | His | Arg | Cys | Gly | Trp | His | Pro | 65 | 70 | 75 | 80 |
| Gly | Thr | Gly | Ser | His | Thr | Gly | Ala | Val | Arg | Arg | Arg | His | Gln | Glu | Ala | 85 | 90 | 95 | |
| Arg | Asp | Gln | Ser | Leu | Leu | Leu | Arg | Arg | Arg | Gly | Arg | Val | Asp | Leu | Asp | 100 | 105 | 110 | |
| Gly | Gly | Gly | Arg | Leu | Arg | Arg | Val | Tyr | Arg | Phe | Gln | Gly | Cys | Leu | Val | 115 | 120 | 125 | |
| Val | Val | Phe | Gly | Gln | His | Leu | Leu | Arg | Pro | Leu | Leu | Ile | Leu | Arg | Val | 130 | 135 | 140 | |
| His | Arg | Glu | Asn | Leu | Val | Ala | Gly | Arg | Arg | Val | Phe | Arg | Val | Lys | Pro | 145 | 150 | 155 | 160 |
| Phe | Glu | Pro | Asp | Tyr | Val | Phe | Ile | Ser | Arg | Met | Phe | Pro | Pro | Ser | Pro | 165 | 170 | 175 | |
| His | Val | Gln | Leu | Arg | Asp | Ile | Leu | Ser | Leu | Leu | Gly | His | Arg | Ser | Ala | 180 | 185 | 190 | |
| Gln | Phe | Gly | His | Val | Glu | Tyr | Pro | Leu | Pro | Leu | Leu | Ile | Gln | Arg | Ser | 195 | 200 | 205 | |
| Leu | Ala | Ser | Gly | Ser | Arg | Ile | Ala | Phe | Pro | Val | Val | Lys | Pro | Pro | Glu | 210 | 215 | 220 | |
| Pro | Leu | Asp | Val | Ala | Leu | Gln | Arg | Gln | Val | Glu | Ser | Val | Pro | Pro | Ile | 225 | 230 | 235 | 240 |
| Arg | Lys | Val | Arg | Glu | Arg | Cys | Ala | Leu | Val | Ala | Arg | Phe | Glu | Leu | Pro | 245 | 250 | 255 | |
| Cys | Arg | Phe | Phe | Glu | Ile | His | Glu | Val | Gly | Phe | Thr | Gly | Arg | Gly | His | | | | |

181

260 265 270
 Pro Arg Arg Ile Gly
 275

(2) INFORMATION FOR SEQ ID NO:186:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 192 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

Arg Val Ala Ala Ser Phe Ile Asp Trp Leu Asp Ser Pro Asp Ser Pro
 1 5 10 15
 Leu Asp Pro Ser Leu Val Ser Ser Leu Leu Asn Ala Val Ser Cys Gly
 20 25 30
 Ala Glu Ser Ser Ala Ser Ser Ser Ala Arg Ser Gly Asn Gly Ser Arg
 35 40 45
 Trp Thr Ser Met Pro Ser Gly Thr Arg Pro Gly Pro Arg Arg Ala Thr
 50 55 60
 Ser Arg Asp Asp Arg Arg Ser Ala Thr Ser Val Ile Pro Ser Arg Arg
 65 70 75 80
 Ser Val Ala Pro Arg Ala Glu Phe Gly Thr Arg Leu Ala Ser His Arg
 85 90 95
 Ala Ser Pro Ser Asn Ala Cys Pro Val Arg Ile Val Thr Ser Ala Ser
 100 105 110
 Gly Arg Pro Ile Ser Ser Pro Pro Ile Val Arg Ser Arg Ser Cys Val
 115 120 125
 Asp Lys Asn Gly Arg Arg Cys Ala Ser Gly Tyr Arg Arg Leu Asn Arg
 130 135 140
 Ala Arg Ser Ser Ser Ile Ala Ala Arg Cys Arg Thr Ile Gly Thr Phe
 145 150 155 160
 Arg Arg Ser Arg Tyr Ser Ala Ser Met Arg Val Ser Thr Asn Ser Pro
 165 170 175
 His Val Thr His Gly Val Ala Pro Gly Val Thr Arg Arg Ile Gly Gly
 180 185 190

(2) INFORMATION FOR SEQ ID NO:187:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 196 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gln | Gln | Arg | Pro | Gln | Met | Cys | Gln | Arg | Val | Ser | Glu | Ile | Gln | Pro | Arg | 1 | 5 | 10 | 15 |
| Thr | Gln | Phe | Phe | Asn | Arg | Cys | Ala | Leu | Pro | His | Tyr | Trp | His | Phe | Pro | 20 | 25 | 30 | |
| Ala | Val | Ala | Val | Phe | Ser | Lys | His | Ala | Ser | Leu | Asp | Glu | Leu | Ala | Pro | 35 | 40 | 45 | |
| Arg | Asn | Pro | Arg | Arg | Ser | Ser | Arg | Arg | Asp | Ala | Glu | Asp | Arg | Arg | Val | 50 | 55 | 60 | |
| Ile | Phe | Ala | Ala | Thr | Leu | Val | Ala | Val | Asp | Pro | Pro | Leu | Arg | Gly | Ala | 65 | 70 | 75 | 80 |
| Gly | Gly | Glu | Ala | Asp | Gln | Leu | Ile | Asp | Leu | Gly | Val | Cys | Arg | Arg | Gln | 85 | 90 | 95 | |
| Ala | Gly | Arg | Val | Arg | Arg | Gly | Gln | Glu | Leu | His | His | Arg | His | Arg | His | 100 | 105 | 110 | |
| Gln | Gly | Ala | Ala | Pro | Asp | Leu | Arg | Arg | Arg | Arg | Arg | His | Arg | Arg | Val | 115 | 120 | 125 | |
| Gln | Gln | His | Arg | Arg | Leu | Gln | Arg | Val | Arg | Gln | Leu | Arg | Arg | Tyr | Val | 130 | 135 | 140 | |
| Gln | Thr | Ala | His | His | Arg | Arg | Phe | Ala | Arg | Thr | Asp | Arg | Val | Arg | His | 145 | 150 | 155 | 160 |
| His | Val | Arg | Gly | Pro | Ser | Asn | His | Arg | Arg | Arg | Arg | Val | Tyr | Arg | Gly | 165 | 170 | 175 | |
| Arg | His | Ser | Gly | Ala | Gly | Gly | Cys | Pro | Ala | Gly | Gly | Ala | Gly | Ser | Val | 180 | 185 | 190 | |
| Gly | Gly | Ser | Ala | | | | | | | | | | | | | 195 | | | |

(2) INFORMATION FOR SEQ ID NO:188:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 311 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

Val Arg Cys Gly Thr Leu Val Pro Val Pro Met Val Glu Phe Leu Thr

183

| | | | |
|---|-----|-----|-----|
| 1 | 5 | 10 | 15 |
| Ser Thr Asn Ala Pro Ser Leu Pro Ser Ala Tyr Ala Glu Val Asp Lys | 20 | 25 | 30 |
| Leu Ile Gly Leu Pro Ala Gly Thr Ala Lys Arg Trp Ile Asn Gly Tyr | 35 | 40 | 45 |
| Glu Arg Gly Gly Lys Asp His Pro Pro Ile Leu Arg Val Thr Pro Gly | 50 | 55 | 60 |
| Ala Thr Pro Trp Val Thr Trp Gly Glu Phe Val Glu Thr Arg Met Leu | 65 | 70 | 75 |
| Ala Glu Tyr Arg Asp Arg Arg Lys Val Pro Ile Val Arg Gln Arg Ala | 85 | 90 | 95 |
| Ala Ile Glu Glu Leu Arg Ala Arg Phe Asn Leu Arg Tyr Pro Leu Ala | 100 | 105 | 110 |
| His Leu Arg Pro Phe Leu Ser Thr His Glu Arg Asp Leu Thr Met Gly | 115 | 120 | 125 |
| Gly Glu Glu Ile Gly Leu Pro Asp Ala Glu Val Thr Ile Arg Thr Gly | 130 | 135 | 140 |
| Glu Ala Leu Leu Gly Asp Ala Arg Trp Leu Ala Ser Leu Val Pro Asn | 145 | 150 | 155 |
| Ser Ala Arg Gly Ala Thr Leu Arg Arg Leu Gly Ile Thr Asp Val Ala | 165 | 170 | 175 |
| Asp Leu Arg Ser Ser Arg Glu Val Ala Arg Arg Gly Pro Gly Arg Val | 180 | 185 | 190 |
| Pro Asp Gly Ile Asp Val His Leu Leu Pro Phe Pro Asp Leu Ala Asp | 195 | 200 | 205 |
| Asp Asp Ala Asp Asp Ser Ala Pro His Glu Thr Ala Phe Lys Arg Leu | 210 | 215 | 220 |
| Leu Thr Asn Asp Gly Ser Asn Gly Glu Ser Gly Glu Ser Ser Gln Ser | 225 | 230 | 235 |
| Ile Asn Asp Ala Ala Thr Arg Tyr Met Thr Asp Glu Tyr Arg Gln Phe | 245 | 250 | 255 |
| Pro Thr Arg Asn Gly Ala Gln Arg Ala Leu His Arg Val Val Thr Leu | 260 | 265 | 270 |
| Leu Ala Ala Gly Arg Pro Val Leu Thr His Cys Phe Ala Gly Lys Asp | 275 | 280 | 285 |
| Arg Thr Gly Phe Val Val Ala Leu Val Leu Glu Ala Val Gly Leu Asp | 290 | 295 | 300 |
| Arg Asp Val Ile Val Ala Asp | 305 | 310 | |

(2) INFORMATION FOR SEQ ID NO:189:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2072 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

| | |
|--|------|
| CTCTGTCCCA TTCCGCACGA GCTGAGCAGC CCAAGGGGCC GTTCGGGCGAA GTCATCGAGG | 60 |
| CATTGSCCGA CCGECTGGCC GGCAAGGGTA AGCAAATCAA CACCACGCTG AACAGCCTGT | 120 |
| CGCAGGCCTT GAAGGCTTTG AATGAGGGGC GCGGCGACTT CTTCGGCGTG GTACGCGAGC | 180 |
| TGGGCTTATT CGTCAACCGG CTACATCAGG ACGACCAACA GTTCGTGCGG TTGAACAAGA | 240 |
| ACCTTGCGGA GTTCACCGAC AGGTGAGCC ACTCGATGC GGACCTGTGG AACGCCATCC | 300 |
| AGCAATTCGA CAGCTTGCTC GCGGTGCGGC GCGCTTCTT CCGCAAGAAC CGCGAGGTGC | 360 |
| TGACGCAVGA CGTCAATAAT CTGCGGACCG TGACCACACG GTTGCTGCGG CCGCATCGT | 420 |
| TGATAGGCTT GGAGACCGTC CTGCACATCT TCCGACGCTT GCGGGCGAAC ATTAACAGC | 480 |
| TTTACCATCC GACACACGGT GCGGTGCTGT CGCTTTCCGC GTTCACGAAT TTGCGCAACC | 540 |
| CGATGGAGTT CATCTGCAGC TCGATTGAGG CCGGTAGCCG GCTCGCTTAT CAGAGTCCG | 600 |
| CCGAACCTCG TCGCGAGTAT CTGGCGCCAG TCCTGATGC GATCAAGTTC AACTACTTTC | 660 |
| CGTTCGCTCT GAACGTGGCC AGCACGCGCT CGACACTGCC TAAAGAGATC GCGTACTCCG | 720 |
| AGCCCCGCTT GCAGCCGCGC AACGGGTACA AGGACACGAC GGTGCGCGGC ATCTGGGTGC | 780 |
| CGGATACGCC GTTGTCACAC CCGAACACGC AGCCCGCTTG GTTGGTGCCA CCGGGGATCC | 840 |
| AAGGGGTTCA GGTGGGACCG ATCAGCGAGG GTTTGCTGAC GCGGGAGTCC CTGGCCCAAC | 900 |
| TCATGCGTGG TCCCGATATC GCGGCTCCGT CGTCAGGCGT GCAAACCGCG CCGGAGCCCC | 960 |
| CGAATGCGTA CGACGAGTAC CCGGTGCTGC CCGCGATCGG TTTACAGGCG CCACAGGTGC | 1020 |
| CGATACCACC GCGGCTCTCT GGGGCGGAGC TAATCGCGGG TCCGCTGCCA CCGGTCTTGG | 1080 |
| CGGCGATCGT GTTCCCAAGA GATCGCCCCG CAGCTTCGGA AAACCTCGAC TACATGGGCC | 1140 |
| TCTTGTTGCT GTCGCGGGGC CTGGCGACCT TCGTGTTCGG GGTGTGATCT AGCCCCGCCC | 1200 |
| GTGGAACGAT GCGCGATCGG CAGGTGTTGA TACCGCGGAT CACCGGCTTG GCGTTGATCG | 1260 |
| CGGCAATCGT CGCACATTCG TGGTACCGCA CAGAACATCC GCTCATASAC ATCGCTTGT | 1320 |
| TCCAGAACCG AGCGGTGCGG CAGGCCAACA TGAGGATGAC GGTGCTCTCC CTCGGGCTGT | 1380 |
| TTGCTCTCTT CTGCTGCTC CCGAGCTACG TCCAGCAAGT GTTGCACCAA TCACCGATGC | 1440 |
| AATCGGGGGT GCATATCATC CCACAGGGCC TCGGTGCCAT GCTGGCGATG CCGATCGCGG | 1500 |
| GAGCGATCAT GCACGAGCGG GGACCGGCCA AGATGCTGCT GGTGCGGATC ATGCTGATCG | 1560 |

```

CTGCGGGGTT GGGCACCTTC GCTTTTGGTG TCGCGGGGCA AGCGGACTAC TTACCCATTC      1620
TGGCGACCGG GCTGGCAATC ATGGSCATGG GCATGGGCGG CTCCATGATE CCACTGTCCG      1680
GGGCGGCGAGT GCAGACCTTG GCGCCACATC AGATGGCTCG CGGTTCGACG CTGATCAGCG      1740
TCAACCAGCA GGTGGGCGGT TCSATAGGGA CCGCACTGAT GTGGGTGCTG CTCACCTAAC      1800
AGTTCAATCA CAGCGAAATC ATCCTACTTG CAAAGAAAGT CCGACTGACC CCAGAGAGTG      1860
GCGCGGGGCG GGGGGGCGCG GTTGACCTTT CCTCGCTACC GCGCCAAACC AACTTCGCGG      1920
CCCAACTGCT GCATGACCTT TCGCAGCGCT ACGCGTGTGT ATTCGTGATA GCGACCGCGC      1980
TAGTGGTCTC GACGCTGATC CCGCGGCGAT TCTTCCGAA ACAGCAGGCT AGTCATCGAA      2040
GAGCACCTTT GGTATCCCA TGACCTCTGC TT                                     2072

```

(2) INFORMATION FOR SEQ ID NO:190:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1923 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

```

TCACCCCGGA GAATCGTTC GTGACGACC TGGACATCGA CTCGCTGTGG ATGGTCGAGA      60
TGGCGGTGCA GACCGAGSAC AAGTACGGCG TCAAGATCCC CGACGAGGAC CTCGCGSCTC      120
TGGTACCGT CGGTGACGTT GTCGGCTACA TCCAGAACTT CGAGGAAGAA AACCCGGAGG      180
CGGCTCAGGC GTTGCGCGCG AAGATTGAGT CGGAGAACCC CGATCCGCA CGAGCAGATC      240
GGTGCCTTTC ACCCACATCG CAAGCTCGAG ACCCGGCTCG TCCTCTTGCA CGCTCAGCCA      300
GGTTGGGCTG TCGCGGCTTT CCAGCAAGTG TTCCGACCAC ACGAGGGGAC CCTCGCGAAA      360
GGTGACTGAT CCGCGGACCA CATAGTCGAT GCGACCTTGG CTGACAATTC CCGCGGCTOC      420
GAGTTGGCGG GGCCCGAATT GCGGCATTGC GTCGAAGGCC AGCGGATCCC GCGCGCGGCC      480
CGGCGTGCTT GGTGTTTTGG GCGCGCGGAT GCGCAGGAGC AGAAGGAGCA TGGCGGCGAT      540
GAACAGCGGC ACGGCAATCA CGACGAGCAG ATTTCCACAG CATACCTCTT CGTACCGCTG      600
CGCGCGGCTT GGTCAATCGG TCGCATATCG ATGGGCGGCT TTAACGTAAC AGCTTTGCGG      660
GGACCGGGGG TCACACGGGG CCAATTGTTC GGCGGGGAAC CCGGCAGGTC TCGGCGCGGG      720
TCACCCAGAG TCACTGGTGC ACCATCCGGG TGTGGGTGAG CGTGCAACTC AAACACACTC      780
AACGGCAACG GTTTCTCAGG TCACAGGCTC AACCTCGACC CGCAATCGCT CGTACGTTTC      840
GACCGCGCGC AGGTGCGGAG TCAGCAGCTT TCGCGCGGCA GCTTTGCGCG TGAAGCGGAC      900

```

```

CAGGGCATCG TAGGTTGCGC CACCGGTSAC ATCGTGCTCG GCGAGGTGCT CGGTCAAGCC      960
GCGATATGAG CAGGCATCCA GTGCCAGSTA GTTGCTGGAG GTGATGTCCG CCAAGTAGGC      1020
GTGGACGGCA ACAGGGGCAA TACGATGCGG CGGTGCTAGC CCGGTCAAGA CCGAATAGGT      1080
TTCCACAGCC GCTGCGCGCA TCAGATGGAC GCCACGGTGG AGCGCGCGCA CGGCGGCTTC      1140
GTGCCCTTGG TGCCAGCTCG CGAATCCGCG AACCAGLACG CTGCTGTCTG GTGCCATCAC      1200
CGCGGTGTGC GATCGAGGGT TTCCCGAAGC ATTTCTGTGG TCAACGGGGG CAGGGGAGGT      1260
TCTGGCCTGG CGACGAGAAC CGAGCCTTCC CGACAGATT CGACACGGT CGGGGCGCGC      1320
TCAATCTGSA TCGGCCCATC GCGCTCGGTG ATCTCCAGCT GTTCCTTCCC GCGCAAGCCA      1380
AGGCGCTGCG GAATCCGCTT GGGAAATCAC AGACCTCTG CGACATCGAT GGTGTTCGC      1440
ATGCTAGGAA ATTTACCATC GCACGTTCCA TAGGCTGTC CTGCGCGGA TGTCGGAGC      1500
ATCCGCTAGC GTATCGAAGC ATTGTTCCG AATGGCTGA GGGAGCTGC GTTCGGGTC      1560
ATGGGTGTG ATCCCGGCTT GAGCCGATGC GGGCTGTGCG TCATCGAGAG TGGGCGTGT      1620
CGGCAGCTCA CCGCGCTGSA TGTGACGCTG GTGCGCACAC CGTCGATGC GGCCTTGGCG      1680
CAGCGCTGT TGGCCATCAG CGATGCGCTC GAGCACTGCG TGGACACCA TCATCCGAG      1740
GTGCTGCTA TCGAACGGGT GTTCTCTCAG CTCACGCTGA CCAGGTGAT GGGCACCGC      1800
CAGGCGGGG GCGTGATCGC CCGCGCGCG GCGAAAGCTG GTGTGACGT GCATTTCCAT      1860
ACCCCGAGCG AGGTCAAGCC GCGCGTCACT GCGAAGGTT CGCGAGCAA GGTTCAGGT      1920
ACC                                               1923

```

(2) INFORMATION FOR SEQ ID NO:191:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1955 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(*) SEQUENCE DESCRIPTION: SEQ ID NO:191:

```

CTGGCGTGCC AGTGTACCC GCGATATGAC GTGCGATTG AATTTGCGG CCCCCCGGA      60
CCCTCGCCA CCCAATCTGG ACCACCGGT CGGTCAATT CCGAAGGTG CCAAGTGGT      120
GCCCAATGTG GTGCTGGTT TCTTGAACGA AGGCTTGCG TATCGGCTG CCTACCCCA      180
AACAACGCCA GTCCAGGAAT CCGGTCCCG GCGGCGGATT CCGAGGGCA TGTGCTAGC      240
GGGGATGGT CAGACGTAAC GGTTCGCTAG GTGGAACCC GCGCCAGGC CCTGCGCGG      300
GCTCATGGCA GCGAATTAG AAAACCGGG ATATTGTCC GCGATTGTCA TACGATGCT      360

```

187

```

AGTGCTTGGT GGTTCGTGTT TAGCCATTGA GTCTGGATGT GTTGAGACCC TGGCCTGGAA      420
GGGGACAAAG TGCTTTTGCC TCTTGCTCCG CCTTTGCCGC CGGACGCGGT GGTGGCGAAA      480
CGGCTGAGT CCGGAATGCT CGGCGGGTTG TCGGTTCCGC TCAGCTGGG AGTGCTGTG      540
CCACCGGATG ATTATGACCA CTGGGCGGCT GCGCGGAGG ACGCGCGCGA TGTGGATGTC      600
CAGGCGGCGG AAGGGGCGGA CCGAGAGGCC GCGGCGATG ACGAGTGGGA TGAGTGGCAG      660
GCCTGGAACG AGTGGGTGGC GGGGAACGCT GAACCCCGCT TTGAGGTGCC ACGGAGTAGC      720
AGCAGCGTGA TTCCGCATTC TCCGCGCGCC GGTAGGAGA GGGGGCGCAG ACTGTCTTA      780
TTTGACCAAT GATCGGCGGT CTGGGTGTTT CCGCGCGCGG CTATGACAACT AGTCATGTG      840
CATGACAAGT TACAGGTATT AGGTCCAGGT TCAACAAGGA GACAGGCAAC ATGGCAACAC      900
GTTTATGAC GGATCGGCAC GCGATGCGGG ACATGGCGGG CCGTTTGGAG GTGCACGCC      960
AGACGGTGA GAGCAGGCT CCGCGGATGT GGGCTTCGCG GCRAAACATC TCGGCGCGG      1020
GCTGGAGTGG CATGGCGGAG GCGACCTCGT TAGAC      1055

```

(2) INFORMATION FOR SEQ ID NO:192:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 359 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:192:

```

CGGCTCGTT GTTGGCATAC TCCGCGCGCG CCGCTCGAC CGCACTGGCC GTGGCTGTG      60
TCCGGCTGA CCACGGGAT CCGCGAAGCA TCCGATCA CCTGCAATC ATCCAGCTG      120
CGCAGCTGCT CACCCAGCCA CCGGCGGTG TCGGAGAGG CCTGCATCAC CTTCGTATAG      180
CCCTCGCGCC CCAGCGCGAG GAGTTGTAG TACTGGCCA CCACCTGCTT ACCGGGACGG      240
GAGAAATTCA GGTTGAGGT CGGCATGTG CCGCGAGGT AGTTGACCG GAAACACAGA      300
TCCTCGGCA GTTCTCGGG CCGCGGCCAC ACGACAAAC CGAGCGCGG ATAGGTGAG      359

```

(2) INFORMATION FOR SEQ ID NO:193:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 350 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(*1) SEQUENCE DESCRIPTION: SEQ ID NO:193:

| | |
|---|-----|
| AACGGGCCCCG TGGGCACCGC TCCTCTAAGG GCTCTCGTTG GTCCGATGAA GTGCTGGAG | 60 |
| GATGCATCTT GGCAGATTCC CGCCAGAGCA AATCAGCCGC TAGTCCTAGT CCGAGTCGCC | 120 |
| CGCAAGTTTC CTCGAATAAC TCCGTACCGG GAGCGCCAAA CCGGCTCTCC TTCGCTAAGC | 180 |
| TGCGCGAACC ACTTGAGGTT CCGGGACTCC TTGACGTCCA GACCGATTCC TTCGAGTGGC | 240 |
| TGATCGGTTT GCAGCGCTGG CGCGAATCCG CCGCCGAGCG GGGTGTATGC AACCCACTCG | 300 |
| GTGGCCTGGA AGAGGTGCTC TACGAGCTGT CTCGATCGA GCACTTCTCC | 360 |

(2) INFORMATION FOR SEQ ID NO:194:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 679 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(*1) SEQUENCE DESCRIPTION: SEQ ID NO:194:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Gln | Pro | Lys | Gly | Pro | Phe | Gly | Glu | Val | Ile | Glu | Ala | Phe | Ala | Asp | 1 | 5 | 10 | 15 |
| Gly | Leu | Ala | Gly | Lys | Gly | Lys | Gln | Ile | Asn | Thr | Thr | Leu | Asn | Ser | Leu | 20 | 25 | 30 | |
| Ser | Gln | Ala | Leu | Asn | Ala | Leu | Asn | Glu | Gly | Arg | Gly | Asp | Phe | Phe | Ala | 35 | 40 | 45 | |
| Val | Val | Arg | Ser | Leu | Ala | Leu | Phe | Val | Asn | Ala | Leu | Sis | Gln | Asp | Asp | 50 | 55 | 60 | |
| Gln | Gln | Phe | Val | Ala | Leu | Asn | Lys | Asn | Leu | Ala | Glu | Phe | Thr | Asp | Arg | 65 | 70 | 75 | 80 |
| Leu | Thr | His | Ser | Asp | Ala | Asp | Leu | Ser | Asn | Ala | Ile | Gln | Gln | Phe | Asp | 85 | 90 | 95 | |
| Ser | Leu | Leu | Ala | Val | Ala | Arg | Pro | Phe | Phe | Ala | Lys | Asn | Arg | Glu | Val | 100 | 105 | 110 | |
| Leu | Thr | Sis | Asp | Val | Asn | Asn | Leu | Ala | Thr | Val | Thr | Thr | Thr | Leu | Leu | 115 | 120 | 125 | |
| Gln | Pro | Asp | Pro | Leu | Asp | Gly | Leu | Glu | Thr | Val | Leu | Sis | Ile | Phe | Pro | 130 | 135 | 140 | |
| Thr | Leu | Ala | Ala | Asn | Ile | Asn | Gln | Leu | Tyr | Sis | Pro | Thr | His | Gly | Gly | 145 | 150 | 155 | 160 |
| Val | Val | Ser | Leu | Ser | Ala | Phe | Thr | Asn | Phe | Ala | Asn | Pro | Met | Glu | Phe | 165 | 170 | 175 | |

189

Ile Cys Ser Ser Ile Gln Ala Gly Ser Arg Leu Gly Tyr Gln Glu Ser
 180 185 190
 Ala Glu Leu Cys Ala Gln Tyr Leu Ala Pro Val Leu Asp Ala Ile Lys
 195 200 205
 Phe Asn Tyr Phe Pro Phe Gly Leu Asn Val Ala Ser Thr Ala Ser Thr
 210 215 220
 Leu Pro Lys Gln Ile Ala Tyr Ser Glu Pro Arg Leu Gln Pro Pro Asn
 225 230 235 240
 Gly Tyr Lys Asp Thr Thr Val Pro Gly Ile Trp Val Pro Asp Thr Pro
 245 250 255
 Leu Ser His Arg Asn Thr Gln Pro Gly Trp Val Val Ala Pro Gly Met
 260 265 270
 Gln Gly Val Gln Val Gly Pro Ile Thr Gln Gly Leu Leu Thr Pro Gln
 275 280 285
 Ser Leu Ala Gln Leu Met Gly Gly Pro Asp Ile Ala Pro Pro Ser Ser
 290 295 300
 Gly Leu Gln Thr Pro Pro Gly Pro Pro Asn Ala Tyr Asp Glu Tyr Pro
 305 310 315 320
 Val Leu Pro Pro Ile Gly Leu Gln Ala Pro Gln Val Pro Ile Pro Pro
 325 330 335
 Pro Pro Pro Gly Pro Asp Val Ile Pro Gly Pro Val Pro Pro Val Leu
 340 345 350
 Ala Ala Ile Val Phe Pro Arg Asp Arg Pro Ala Ala Ser Gln Asn Phe
 355 360 365
 Asp Tyr Met Gly Leu Leu Leu Leu Ser Pro Gly Leu Ala Thr Phe Leu
 370 375 380
 Phe Gly Val Ser Ser Ser Pro Ala Arg Gly Thr Met Ala Asp Arg His
 385 390 395 400
 Val Leu Ile Pro Ala Ile Thr Gly Leu Ala Leu Ile Ala Ala Phe Val
 405 410 415
 Ala His Ser Trp Tyr Arg Thr Gln His Pro Leu Ile Asp Met Arg Leu
 420 425 430
 Phe Gln Asn Arg Ala Val Ala Gln Ala Asn Met Thr Met Thr Val Leu
 435 440 445
 Ser Leu Gly Leu Phe Gly Ser Phe Leu Leu Leu Pro Ser Tyr Leu Gln
 450 455 460
 Gln Val Leu His Gln Ser Pro Met Gln Ser Gly Val His Ile Ile Pro
 465 470 475 480
 Gln Gly Leu Gly Ala Met Leu Ala Met Pro Ile Ala Gly Ala Met Met
 485 490 495
 Asp Arg Arg Gly Pro Ala Lys Ile Val Leu Val Gly Ile Met Leu Ile
 500 505 510

190

```

Ala Ala Gly Leu Gly Thr Phe Ala Phe Gly Val Ala Arg Gln Ala Asp
515                               520                               525

Tyr Leu Pro Ile Leu Pro Thr Gly Leu Ala Ile Met Gly Met Gly Met
530                               535                               540

Gly Cys Ser Met Met Pro Leu Ser Gly Ala Ala Val Gln Thr Leu Ala
545                               550                               555                               560

Pro His Gln Ile Ala Arg Gly Ser Thr Leu Ile Ser Val Asn Gln Gln
565                               570                               575

Val Gly Gly Ser Ile Gly Thr Ala Leu Met Ser Val Leu Leu Thr Tyr
580                               585                               590

Gln Phe Asn His Ser Glu Ile Ile Ala Thr Ala Lys Lys Val Ala Leu
595                               600                               605

Thr Pro Glu Ser Gly Ala Gly Arg Gly Ala Ala Val Asp Pro Ser Ser
610                               615                               620

Leu Pro Arg Gln Thr Asn Phe Ala Ala Gln Leu Leu His Asp Leu Ser
625                               630                               635                               640

His Ala Tyr Ala Val Val Phe Val Ile Ala Thr Ala Leu Val Val Ser
645                               650                               655

Thr Leu Ile Pro Ala Ala Phe Leu Pro Lys Gln Gln Ala Ser His Arg
660                               665                               670

Arg Ala Pro Leu Leu Ser Ala
675

```

(2) INFORMATION FOR SEQ ID NO:195:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 120 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

```

Thr Pro Glu Lys Ser Phe Val Asp Asp Leu Asp Ile Asp Ser Leu Ser
1           5           10           15

Met Val Glu Ile Ala Val Gln Thr Glu Asp Lys Tyr Gly Val Lys Ile
20           25           30

Pro Asp Glu Asp Leu Ala Gly Leu Arg Thr Val Gly Asp Val Val Ala
35           40           45

Tyr Ile Gln Lys Leu Glu Glu Glu Asn Pro Glu Ala Ala Gln Ala Leu
50           55           60

Arg Ala Lys Ile Gln Ser Glu Asn Pro Asp Ala Ala Arg Ala Asp Arg
65           70           75           80

```

191

```

Cys Val Ser Pro Thr Ser Gln Ala Arg Asp Ala Arg Arg Pro Leu Ala
      85                      90                      95
Arg Ser Ala Arg Leu Ala Cys Arg Arg Leu Pro Ala Ser Val Pro Thr
      100                      105                      110
Thr Arg Arg Asp Pro Arg Glu Arg
      115                      120

```

(2) INFORMATION FOR SEQ ID NO:196:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 89 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

```

Leu Ala Cys Gln Cys His Arg Arg Tyr Asp Val Gly Ile Gln Phe Arg
1          5          10          15
Gly Pro Ala Gly Pro Val Ala Thr Gln Ser Gly Pro Pro Gly Pro Ser
      20          25          30
Ile Ala Gln Gly Arg Gln Val Arg Ala Gln Cys Gly Ala Gly Phe Leu
      35          40          45
Gln Arg Arg Pro Ala Val Ser Gly Ala Leu Pro Pro Asn Asn Ala Ser
      50          55          60
Pro Gly Ile Arg Ser Arg Ala Ala Asp Ser Gln Arg His Leu Leu Ala
      65          70          75          80
Gly Asp Gly Ser Asp Val Thr Val Gly
      85

```

(2) INFORMATION FOR SEQ ID NO:197:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 119 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

```

Ala Ser Leu Leu Ala Tyr Ser Ala Ala Ala Ala Ser Thr Ala Leu Ala
1          5          10          15
Val Ala Cys Val Arg Ala Asp His Arg Asp Arg Arg Thr Ile Arg Asp
      20          25          30

```


192

His Leu Ala Met Ile His Leu Ala Glu Leu Val Thr Gln Pro Pro Gly
 35 40 45
 Gly Val Arg Gln Arg Leu His His Leu Gly Ile Ala Val Ala Pro Gln
 50 55 60
 Pro Gln Glu Val Val Val Leu Ala His His Leu Val Thr Gly Thr Gly
 65 70 75 80
 Glu Val Gln Gly Glu Gly Arg His Val Ala Ala Glu Val Val Asp Pro
 85 90 95
 Glu Asn Gln Ile Leu Arg Gln Val Leu Gly Pro Ala Pro His Asp Lys
 100 105 110
 Pro Asp Ala Gly Ile Gly Gln
 115

(2) INFORMATION FOR SEQ ID NO:198:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 116 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

Arg Ala Arg Gly His Arg Ser Ser Lys Gly Ser Arg Trp Ser His Gln
 1 5 10 15
 Val Leu Glu Gly Cys Ile Leu Ala Asp Ser Arg Gln Ser Lys Thr Ala
 20 25 30
 Ala Ser Pro Ser Pro Ser Arg Pro Gln Ser Ser Ser Asn Asn Ser Val
 35 40 45
 Pro Gly Ala Pro Asn Arg Val Ser Phe Ala Lys Leu Arg Glu Pro Leu
 50 55 60
 Glu Val Pro Gly Leu Leu Asp Val Gln Thr Asp Ser Phe Glu Trp Leu
 65 70 75 80
 Ile Gly Ser Pro Arg Trp Arg Glu Ser Ala Ala Glu Arg Gly Asp Val
 85 90 95
 Asn Pro Val Gly Gly Leu Glu Glu Val Leu Tyr Glu Leu Ser Pro Ile
 100 105 110
 Glu Asp Phe Ser
 115

(2) INFORMATION FOR SEQ ID NO:199:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 811 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:199:

| | |
|--|-----|
| TCCTACGCAG CAATCGCTTT GGTGACAGAT GTGGATGCCG GCGTCGCTGC TGGCGATGCC | 60 |
| GTGAAAGCCG CCGACGTGTT CGCCGCRATC GGGGAGABCA TCGAAGTGGT CAAAAGGCTG | 120 |
| GTGCGGSCCG CCATCCATCG GGTCCCGGAC GAGCGCACGT GCACGCACTG TCAACACCCAC | 180 |
| GCCGGTGTTC CGTTGCCGTT CGAGCTGCCA TGAGGGTGCT GCTGACCGGC GCGGCCCGCT | 240 |
| TCATCGGCTC GCGCGTGGAT GCGGCGTTAC GGGCTGCGGG TCAAGACGTC GTGGGCTGCG | 300 |
| ACGCGCTGCT GCGCGCGCGC CACGGGCGAA ACCCGCTGCT GCCACCGGGC TGGCAGCGCG | 360 |
| TGCACGTGCG CCACGCCAGC GCGCTGGCGC CGTTGTGGC CGGTGTGGAT CTGGTGTGTC | 420 |
| ACCAGGCGGC CATGGTGGCT GCGCGCGTCA ACCCGCGCGA CGCACCGGCG TATGCGCGGC | 480 |
| ACAACGATTT CGCCACCACG GTGCTGCTGC CGCAGATGTT CCGCGCGCGG GTCCGCGCTT | 540 |
| TGGTGTGCGC GTGCTGATG GTGCTTTACG GCGCGCGCGC CTATGACTGT CCGCAGCATG | 600 |
| GACCGGTGCA CCGGCTGCGC GCGCGCGGAG CCGACCTGCA CAATGCGGTC TTCGAGCGCC | 660 |
| GTTGCGCGCG GTGCGCGGAG CAGTGCATCT GCGAATGCTT CGACGAGAT GCGCGCTTGC | 720 |
| GCGCGCGCAG CCTGTACGCG GCAGCAAGAC CGCGCAGGAG CACTACGCGC TGGCGTGGTC | 780 |
| GGAAACGATT GCGCGTTCG TGGTGGCGTT G | 811 |

(2) INFORMATION FOR SEQ ID NO:200:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 966 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:200:

| | |
|---|-----|
| GTCCCGCGAT GTGGCGAGC ATGACTTTCC GTCACACCGG CGTAGTAGTC GAGATATCC | 60 |
| GACTTTGTGG TCCCGGTGGC GGGATAGGCG ACCTGTGCGC GTTGGTCAGC GTACCGGTT | 120 |
| GCTCGGACGC CGAACCCTG CTTTCAACGT AGCTGTGCG TCAACACAGT CCGGAGCCTA | 180 |
| ACGTACCGGT CAAATATCCG GTGGAATTTG GCGGTGAGT TCCGCTCGCG GACAATCAAG | 240 |
| GCACTACTAC TTACATGCGA GCGATTTGGA CGGTTTCGAT CGCCTTGGCG CTGGTGAACG | 300 |
| TGCGCGTCAA GGTGTACAGC GCTACCGCAG ACCACGACAT CAGGTTCCAC CAGGTGCACG | 360 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| CCAGGACAAA | CGGACGCATC | CGGTACAGAC | GCSTCTGCGA | GGCCTGTGEC | GAGGTGGTCC | 420 |
| ACTACCGCGA | TCTTGCCCGG | GCCTACGAGT | CGGCGAGCGG | CCAAATGGTG | GCATCACC | 480 |
| ACGACGACAT | CGCCAGCTTG | CCTGAAGAAC | GCAGCCGCGA | GATCGAGGTG | TTGCASTTCC | 540 |
| TCCCCCGCCG | CGACGTGGAC | CGGATGATGT | TGAGCCGCGG | CTACTTTTTG | GAGCCTGATT | 600 |
| CGAAGTCGTC | GAATTCGTAT | GTGCTGCTGG | CTAAGACACT | CGGCGAGACC | GACCGCATCC | 660 |
| CGATCGTGGA | TGCCCCCACC | GGCCGTGAAT | GCAGGAAAAA | TAAGAGCCCG | TATCCACAA | 720 |
| TGCGCGTCCA | GCTCGGCTAC | CACAAACGGT | AGAACGATCG | AGACATTCGC | GAGCTGAAGT | 780 |
| GGGGCGCTAT | AGAAGCCGCT | CTGCGCGATT | ATCAAAACGA | AAATACGCTT | ACTCATGCCA | 840 |
| TGCGCGCTGC | TCACCGCATG | CGACGTTTTT | GGCAGCGCTC | ACCGCGTGGC | CGCGGACCTC | 900 |
| AAGTGGGCAT | GCATCCGACC | CGTCCCCGGA | AACCGGTTCC | GGCGGGTGCG | CTCATCGCTT | 960 |
| CATCCY | | | | | | 966 |

(2) INFORMATION FOR SEQ ID NO:201:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2367 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| CCCCACCGCC | GGCAATACCG | CCAGCGGCAC | CGTTACCGCC | GTTTGGCGCG | TTGCCCCCGT | 60 |
| TGCGCGCCGT | CCCGCGCGCC | CCCGCGATGG | AGTTCTCATC | GCCAAAAGTA | CTGGCGTTGC | 120 |
| CACCGGAGCC | GGCGTTGCGG | CGGTACCGCG | CAGCCCCGCG | GACTCCACCG | GGCCACCGCA | 180 |
| CTCGCGCGCT | GCCACCGTTG | CGCGCGTTGC | CGATCAACAT | GGCGCTGGCG | CCACCGTTGC | 240 |
| CACCCACGCG | ACCGGCTGCG | CCACCGCGCG | CGACACCAAG | CGAGCTGCGG | CCGAGCGCAC | 300 |
| CATCACCACC | TACGCGACCG | ACCGCGCAGA | CACGAGCGAC | CGGCTCTTCC | TGAAACGTCG | 360 |
| CGGTGCTACC | ACCGCGCGCG | TTACCGCGCA | CCCCACCGCG | AACCGCGCGG | CGGCGATGCC | 420 |
| CGCGCGCGCG | GGCGTTGCGG | CGGTGCGCG | CGTTGCGCGA | CAACACCGCG | CCCGCGCGCG | 480 |
| CGTTGCGCGG | CGCGCGCGCG | GTCCCGCGCG | CGCGCGCGAC | GCCAAGCGCG | CTCGCGCGCT | 540 |
| TGCGCGCATC | ACCACCGTTG | CCCGCGACCA | CATCGCGTTT | TGCGTGGCGG | TCTGCGCTGT | 600 |
| CAACCGTGG | GATGCGAGCG | TTGCGCGCGC | TTCCCCCGGG | CCCGCGCGTG | GGCGCGTGC | 660 |
| CACCGATACC | ACCGCGCGCA | CGCGCGCGAC | CGTTGCGCGC | ATCACCGAAT | AGCAACCGCG | 720 |
| CGCGCGCAC | ATTGCGCGCA | GCTCGCGCTG | CGCGCGCGTC | GGCGCGCGAG | GGCGCGCTGG | 780 |

| | | | | | | |
|-------------|------------|------------|------------|------------|------------|------|
| CAGCCCGCTT | ACCACCGAAA | CCGCCGCTAC | CACCGGTAGA | GCTGGCAGTG | GCGATGTGTA | 840 |
| CGAAAGCGCC | GCCTCGGCG | CCGCCGCTAC | CACCCCGACT | GCCTCGGCT | ACACCGTCGG | 900 |
| ACCGTTTSCC | ACCATCACCG | CCAAAGGCGC | TGCAATGTC | GCCTGCGCG | ACTCCGCCST | 960 |
| CGCCGCTCTT | GCCTCGGCG | CCACCGGCG | CGCGGTACC | GCGTCACCA | CCGCGACCCG | 1020 |
| CGCTGGGCTT | GCCCGAGGCT | GCCTCGGCG | TGCGACGCTC | GCGCGGCTG | CCACCGGTCG | 1080 |
| GCCTGCGGCG | AGTGCCATGG | CCGCCGCTGC | CGCGGTGCG | GCGGTTTGA | TGACCGATGC | 1140 |
| CGGACACATC | TGCGGCGCTG | TCCCGGCTGC | TGCGCGGCG | GCGCGGCTG | GGATTGACCC | 1200 |
| CGTTTGGCCC | GGCGAGGCG | GCGCGGCGG | TACCGCGCG | GCGCGATGG | CGAAGAGCC | 1260 |
| CGCGGTTGCC | GCGTTACCG | CCCGCACCCG | CGATGCTGCG | GGCGAGGCTG | GTGCGCGCGA | 1320 |
| CACCGCGCTT | GCGCGGCTG | CCCGACAGCG | ACCGCGGCTT | CCCGCGGCG | CGCGCGGCGG | 1380 |
| CGCGGCTACG | ACCGCGGCG | CGCTGCGCG | CGTTGCGGAT | CGCGCGGCG | GCGCGCGCG | 1440 |
| TGCGCGGCTT | GTGACGAGCG | CGCGCGGCG | CGCGGTTGCC | ACCGTTGCCA | BACAGCAAGC | 1500 |
| CGCGCGGCGG | GCGAGGCTGC | CGCGGTTGCC | TCCCGTGGCG | GCGGTTTCCG | ATCAACGCGC | 1560 |
| GCGCGAAAG | CGCGTGGCTG | GCGCGATTCA | CGCGACCGCG | CAGACTCGCG | TCAACAGCGG | 1620 |
| CTTCAGTGCT | GGCATACCGA | CCCGCGGCGG | CATCGACGCG | CTCGACAAAC | TGCTCGTGAA | 1680 |
| AGCGTGCCAC | CTGTAGGCTG | AGCGGCTGAT | ACTGCGGAGC | ATGGGCGCGG | AACAACCGCG | 1740 |
| CAATCGCGCG | CGACACTTCA | TGCGCGGCGG | CGCGGCGCG | TTCGGTCGTC | GGGATCGCGG | 1800 |
| CGCGCGGCTT | AGCGCGGCTG | ACCTCGGAGC | CAATAGTCTA | TAAATCGAAA | GCGCGAGTTG | 1860 |
| CCAGCAGGCTG | CGCGGTCGCG | ATCGCGGAGG | ACCGCTCGCA | CTTCGCGATA | CCCGATATCG | 1920 |
| CGCGACCGTG | TCCCGAGCGG | CCAGGTCAGC | TTTGGTCGCT | GCGTGGCGCG | CCTGACTATG | 1980 |
| GCGCGGAGCG | CGCTCGTTCT | GATTCGCGCG | GCGCGGCGCG | TTGTTGCGCG | AGTTGAGAGC | 2040 |
| GGGAGGACAG | GCGGAGCTTG | GTGTAGAGCT | GCGTCAAGTG | GGAATGACAG | GTCCGCGCGG | 2100 |
| AGATGAATAG | GCGGAGCGCG | ATCTGCTTGT | TGCTGAGTCC | CTCAGCGAGC | AGTAGAGCGA | 2160 |
| CCTCAAGCTC | TGTCGCTGTC | AACGCGCGCG | AGCGACTTCT | CGCGCGTTTC | CGTGACCGCG | 2220 |
| GCGCTCGTTG | CGCGTACCGG | ATCGCGCTCA | CGATCGATAA | CGCGTTCTCT | TGCGCGCGCG | 2280 |
| CATCGTCGAA | CTCGCTGTCA | CCCATGGATT | TTCGAAGGCT | GCGTAGCGAG | GAGTTACAGC | 2340 |
| CCCGCTCGTA | GATCCCGAAG | CGGACCG | | | | 2367 |

(2) INFORMATION FOR SEQ ID NO:202:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 376 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:202:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gln | Pro | Ala | Gly | Ala | Thr | Ile | Ala | Ala | Ser | Ser | Pro | Cys | Ala | Thr | Val | 1 | 5 | 10 | 15 |
| Gly | Ala | Gly | Gly | Gly | Thr | Gly | Ser | Pro | Val | Thr | Thr | Gln | Thr | Ala | Ala | 20 | 25 | 30 | |
| Thr | Thr | Gly | Arg | Gly | Gly | Ser | Gly | Asp | Val | Tyr | Gln | Ser | Ala | Ala | Ser | 35 | 40 | 45 | |
| Gly | Ala | Ala | Ala | Thr | Thr | Pro | Thr | Ala | Gly | Gly | Tyr | Thr | Val | Gly | Pro | 50 | 55 | 60 | |
| Val | Ala | Thr | Ile | Thr | Ala | Lys | Gly | Ala | Arg | Asn | Val | Ala | Leu | Arg | Asp | 65 | 70 | 75 | 80 |
| Ser | Ala | Val | Ala | Ala | Val | Ala | Ala | Ala | Ala | Thr | Gly | Ser | Gly | Gly | Thr | 85 | 90 | 95 | |
| Ala | Val | Thr | Thr | Gly | Thr | Ala | Gly | Gly | Leu | Ala | Arg | Ala | Cys | Arg | Arg | 100 | 105 | 110 | |
| Gly | Gly | Thr | Val | Ala | Ala | Gly | Ala | Thr | Gly | Arg | Arg | Ala | Gly | Ser | Ala | 115 | 120 | 125 | |
| Met | Ala | Ala | Arg | Ala | Ala | Val | Ala | Ala | Gly | Leu | Ile | Thr | Asp | Ala | Gly | 130 | 135 | 140 | |
| His | Ile | Cys | Arg | Ala | Val | Pro | Gly | Ala | Gly | Arg | Gly | Ala | Gly | Arg | Gly | 145 | 150 | 155 | 160 |
| Ile | Asp | Pro | Val | Cys | Pro | Gly | Gln | Ala | Gly | Ala | Ala | Gly | Thr | Thr | Gly | 165 | 170 | 175 | |
| Ala | Ala | Met | Ala | Gln | Gln | Pro | Gly | Val | Ala | Ala | Val | Thr | Ala | Arg | Thr | 180 | 185 | 190 | |
| Pro | Asp | Ala | Cys | Gly | His | Ala | Gly | Ala | Ala | Asp | Thr | Ala | Val | Ala | Ala | 195 | 200 | 205 | |
| Val | Ala | Pro | Gln | Pro | Pro | Pro | Val | Pro | Thr | Gly | Thr | Ala | Gly | Arg | Ala | 210 | 215 | 220 | |
| Gly | Thr | Thr | Gly | Pro | Ala | Val | Ala | Ala | Val | Ala | Asp | Gln | Pro | Gly | Arg | 225 | 230 | 235 | 240 |
| Ala | Ser | Ala | Ala | Ala | Gly | Leu | Thr | Gln | Pro | Ala | Ser | Arg | Ala | Val | Ala | 245 | 250 | 255 | |
| Thr | Val | Ala | Lys | Gln | Gln | Pro | Ala | Gly | Arg | Ala | Arg | Leu | Pro | Gly | Cys | 260 | 265 | 270 | |
| Arg | Pro | Val | Gly | Ala | Val | Ser | Asp | Gln | Arg | Ala | Pro | Gln | Lys | Arg | Leu | 275 | 280 | 285 | |
| Gly | Gly | Arg | Ile | His | Arg | Thr | Gln | Gln | Thr | Pro | Leu | Asn | Ser | Gly | Phe | 290 | 295 | 300 | |

Ser Ala Gly Ile Pro Thr Arg Gly Arg Ser Gln Arg Leu His Lys Leu
 305 310 315 320
 Leu Val Lys Arg Cys His Leu Tyr Ala Gln Arg Leu Ile Leu Pro Ser
 325 330 335
 Met Gly Pro Glu Gln Pro Arg Asn Arg Arg Arg His Phe Ile Gly Ser
 340 345 350
 Arg Ser His His Phe Arg Arg Arg Asp Arg Arg Gly Arg Ile Ser Arg
 355 360 365
 Ala His Leu Arg Thr Asn Ser Arg
 370 375

(2) INFORMATION FOR SEQ ID NO:203:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2852 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

GGCCAAAACG CCGCGGCEAT CCGCGCCACC GAGGCGGCTT ACGACCATAT GTGGGCCCCAG 60
 GACGTGGCGG CGATCTTTGG CTACCATGCC GGGGCTTCGG CGGCCCTCTC GCGCTTGACA 120
 CCGTTCGGCC AGGCGCTGCC GACCGTGGCC GCGCGCGCTG CGCTGGTCAG CCGCGCCGCG 180
 GCTCAGGTCA CCACGGCGGT CTTCGGCAAC CTGGGCTTGG CGACGCTCCG CGAGGGGCAAC 240
 GTCCGCAACG GTAATGTCCG GAACTTCAAT CTGGGCTCGG CCAACATCGG CAACGGCAAC 300
 ATCGGCACCG GCAACATCGG CAGCTCCAAC ATCGGGTTTG GCAACGTGGG TCCTGGGTTG 360
 ACCGCAGCUC TGAACAACAT CCGTTTCGCG AACACCGGCA GCAACAACAT CCGGTTTGCG 420
 AACACCGGCA GCAACAACAT CCGGTTTCGG AATACCGGAG ACGGCAACCG AGGTATCGGG 480
 CTCACGGGTA GCGGTTTGTG GCGGTTTCGG GCGCTGAACT CGGGCACCGG CACCATCGGT 540
 CTGTTCAACT CCGGCACCGG AAACSTCGGC ATCGGCACCT CCGGTACCGG GAACTGGGGC 600
 ATTGGCAACT CCGGCACACG CTACAACACC GGTFTTGGCA ACTCGGCGCA CCGCAACACG 660
 GCTTCTTCA ACTCGGCAAT AGCCACACCG GCGCTGGGCA ACGCGGCAAC CTBCAACACC 720
 GGTAGCTTCA ACCCGGGGAA CAGCAATACT GCGGCTTCA ACATGGGCTA GTACAACACG 780
 GGCTAATGTA ACAGCGGCAA CTACAACACC GCGTTGGCAA ACTCCGGCAA TGTCAACACC 840
 GCGGCTTCA TTACTGGCAA CTCAACAAC GCGTCTTGT GCGCGGCGGA CCACCAAGGC 900
 CTGATTTTGG GGAGCGCGCG CTCTTCAAC TCGACCACTG CCGGCTCGTC GGGATTCTTC 960

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| AACAGCGGTG | CCGCTAGGCG | CTCCGGCTTC | CTGAACCTCC | GTGCCAACAA | TTCTGGCTTC | 1020 |
| TTCAACTCTT | CGTCGGGGGC | CATCGGTAAC | TCCGGGCTGG | CAACCGGGG | CGTGCTGGTA | 1080 |
| TCCGGCGTGA | TCAACTCGGG | CAACACCGTA | TGGGTTTTGT | TCAACATGAG | CCTGGTGGCC | 1140 |
| ATCACAAGCG | CGGCTTGAT | CTCGGGCTTC | TTCAACACCG | GAAGCAACAT | GTCCGGATTT | 1200 |
| TTGGGTGGGC | CACCGGTCTT | CAATCTCGGC | CTGGCAAAAC | GGGGGTCTCT | GAACATTTCT | 1260 |
| GGCAACGCCA | ACATCGGCA | TTACAACATT | CTCGGCAGCG | GAAAGCTCGG | TGACTTCAAC | 1320 |
| ATCCTTGGCA | GGGCAACCT | CGGCAGCCAA | AACATCTTGG | GCAGCGGCAA | CGTCGGCAGC | 1380 |
| TTCAATATCG | GCAGTGGAAA | CATCGGAGTA | TTCAATCTCG | GTTCGGGAG | CCTGGGAAAC | 1440 |
| TACAACATCG | GATCCGGAAA | CCTCGGGATC | TACAACATCG | GTTTTGGAAA | CGTCGGCGAC | 1500 |
| TACAACCTCG | GCTTCGGGAA | CGCGGGCGAC | TTCAACCAAG | GCTTTGGCAA | CACCGGCAAC | 1560 |
| AACAACATCG | GTTTCGGCAA | CACCGGCAAC | AACAACATCG | GCATCGGGCT | GTCCGGCGAC | 1620 |
| AACCAGCAGG | GCTTCATAT | TGCTAGCGGC | TGGAACCTCG | GCACCGGCAA | CAGCGGCTCG | 1680 |
| TTCAATTCCG | GCACCAATAA | CGTTGGCATC | TTCAACCGCG | GCACCGGAAA | CGTCGGCATC | 1740 |
| GCAAACTCGG | GCACCGGAAA | CTGGGGTATC | CGGAACCCGG | GTACCGACAA | TACCGGCATC | 1800 |
| CTCAATGCTG | GCAGCTACAA | CACGGGCTTC | CTCAACGGCG | GCGACTTCAA | CACGGGCTTC | 1860 |
| TACAACACCG | GCAGCTACAA | CACCGGCGGC | TTCAACCTCG | GTAACACCAA | CACCGGCAAC | 1920 |
| TTCAACGTGG | GTGACACCAA | TACCGGCAGC | TATAACCCGG | GTGACACCAA | CACCGGCTTC | 1980 |
| TTCAATCCCG | GCAGCTACAA | TACCGGCTTC | TTGACACCGG | GGGACTTCAA | CAATGGCTTC | 2040 |
| TTGGTGGCGG | GGATATACCA | GGGCTAGATT | GCCATCGATC | TCTCGGTAC | CACCTCCATT | 2100 |
| ATCCCCATAA | ACGAGCAGAT | GCTCATTGAC | GTACACAACG | TAATGACCTT | CGCGGGCAAC | 2160 |
| ATGATCAGCG | TCACCGAGGC | CTCGACCGTT | TTCCCCAAA | CCTTCTATCT | GAGCGGTTTG | 2220 |
| TTCTTCTTCC | GGCGGCTCAA | TCTCAGGCGA | TCCAGGCTGA | CCCTTCGGAC | GATCACCTTC | 2280 |
| ACCATCGGGG | GACCGACGGT | GACCGTCCCG | ATCAGCATTG | TCGGTGTCT | GGAGAGGCCG | 2340 |
| ACGATTACCT | TCCTCAAGAT | CGATCCGGCG | CCGGGCATCG | GAAATTCGAC | CACCAACCCC | 2400 |
| TGCTCCGGCT | TCTTCAACTC | GGGCACGGGT | GGCAGATCTG | GCTTCCAAAA | CGTCGGGCGC | 2460 |
| GGCAGTTCA | GGCTCTGGAA | CAGTGGTTTG | AGCAGCGCGA | TAGGGAAATC | GGGTTTCCAG | 2520 |
| AACCTCGGGT | CGCTGCAGTC | AGGCTGGGGG | AACCTGGGCA | ACTCCGTATC | GGGCTTTTTC | 2580 |
| AACACCACTA | CGGTGAACCT | CTCCACGCGG | GCCAAATCTT | CGGGCTTCGA | CAACATCGGC | 2640 |
| ACCAACCTGT | CUGGCGTCTT | CGGCGGTCCG | ACCGGGACGA | TTTTCAACGC | GGGCTTGGCC | 2700 |
| AACCTGGGGC | AGTTGAACAT | CGGCAGGGGC | TGCTGCCGAA | TTGGGCACGA | GTTAGATACG | 2760 |
| GTTTCAACAA | TCAATCCGC | GTTTTGCGGC | AGTGCATCAG | ACGAATCGAA | CCCGGGAGGC | 2820 |

GTATGGCGAAT AAACCGAATG GCGGGCCTGTC AT

2852

(2) INFORMATION FOR SEQ ID NO:204:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 943 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

Gly Gln Asn Ala Pro Ala Ile Ala Ala Thr Gln Ala Ala Tyr Asp Gln
 1 5 10 15
 Met Trp Ala Gln Asp Val Ala Ala Met Phe Gly Tyr His Ala Gly Ala
 20 25 30
 Ser Ala Ala Val Ser Ala Leu Thr Pro Phe Gly Gln Ala Leu Pro Thr
 35 40 45
 Val Ala Gly Gly Gly Ala Leu Val Ser Ala Ala Ala Ala Gln Val Thr
 50 55 60
 Thr Arg Val Phe Arg Asn Leu Gly Leu Ala Asn Val Arg Gln Gly Asn
 65 70 75 80
 Val Arg Asn Gly Asn Val Arg Asn Phe Asn Leu Gly Ser Ala Asn Ile
 85 90 95
 Gly Asn Gly Asn Ile Gly Ser Gly Asn Ile Gly Ser Ser Asn Ile Gly
 100 105 110
 Phe Gly Asn Val Gly Pro Gly Leu Thr Ala Ala Leu Asn Asn Ile Gly
 115 120 125
 Phe Gly Asn Thr Gly Ser Asn Asn Ile Gly Phe Gly Asn Thr Gly Ser
 130 135 140
 Asn Asn Ile Gly Phe Gly Asn Thr Gly Asp Gly Asn Arg Gly Ile Gly
 145 150 155 160
 Leu Thr Gly Ser Gly Leu Leu Gly Phe Gly Gly Leu Asn Ser Gly Thr
 165 170 175
 Gly Asn Ile Gly Leu Phe Asn Ser Gly Thr Gly Asn Val Gly Ile Gly
 180 185 190
 Asn Ser Gly Thr Gly Asn Trp Gly Ile Gly Asn Ser Gly Asn Ser Tyr
 195 200 205
 Asn Thr Gly Phe Gly Asn Ser Gly Asp Ala Asn Thr Gly Phe Phe Asn
 210 215 220
 Ser Gly Ile Ala Asn Thr Gly Val Gly Asn Ala Gly Asn Tyr Asn Thr
 225 230 235 240
 Gly Ser Tyr Asn Pro Gly Asn Ser Asn Thr Gly Gly Phe Asn Met Gly

200

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | | 245 | | | | | | 250 | | | | 255 | |
| Gln | Tyr | Asn | Thr | Gly | Tyr | Leu | Asn | Ser | Gly | Asn | Tyr | Asn | Thr | Gly | Leu |
| | | | 260 | | | | | 265 | | | | | 270 | | |
| Ala | Asn | Ser | Gly | Asn | Val | Asn | Thr | Gly | Ala | Phe | Ile | Thr | Gly | Asn | Phe |
| | | 275 | | | | | 280 | | | | | 285 | | | |
| Asn | Asn | Gly | Phe | Leu | Trp | Arg | Gly | Asp | Ala | Gln | Gly | Leu | Ile | Phe | Gly |
| | | 290 | | | | 295 | | | | | 300 | | | | |
| Ser | Pro | Gly | Phe | Phe | Asn | Ser | Thr | Ser | Ala | Pro | Ser | Ser | Gly | Phe | Phe |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 |
| Asn | Ser | Gly | Ala | Gly | Ser | Ala | Ser | Gly | Phe | Leu | Asn | Ser | Gly | Ala | Asn |
| | | | | 325 | | | | | 330 | | | | | 335 | |
| Asn | Ser | Gly | Phe | Phe | Asn | Ser | Ser | Ser | Gly | Ala | Ile | Gly | Asn | Ser | Gly |
| | | | 340 | | | | | 345 | | | | | 350 | | |
| Leu | Ala | Asn | Ala | Gly | Val | Leu | Val | Ser | Gly | Val | Ile | Asn | Ser | Gly | Asn |
| | | 355 | | | | | 360 | | | | | 365 | | | |
| Thr | Val | Ser | Gly | Leu | Phe | Asn | Met | Ser | Leu | Val | Ala | Ile | Thr | Thr | Pro |
| | 370 | | | | | 375 | | | | | 380 | | | | |
| Ala | Leu | Ile | Ser | Gly | Phe | Phe | Asn | Thr | Gly | Ser | Asn | Met | Ser | Gly | Phe |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 |
| Phe | Gly | Gly | Pro | Pro | Val | Phe | Asn | Leu | Gly | Leu | Ala | Asn | Arg | Gly | Val |
| | | | | 405 | | | | | 410 | | | | | 415 | |
| Val | Asn | Ile | Leu | Gly | Asn | Ala | Asn | Ile | Gly | Asn | Tyr | Asn | Ile | Leu | Gly |
| | | | 420 | | | | | 425 | | | | | 430 | | |
| Ser | Gly | Asn | Val | Gly | Asp | Phe | Asn | Ile | Leu | Gly | Ser | Gly | Asn | Leu | Gly |
| | | 435 | | | | | 440 | | | | | 445 | | | |
| Ser | Gln | Asn | Ile | Leu | Gly | Ser | Gly | Asn | Val | Gly | Ser | Phe | Asn | Ile | Gly |
| | 450 | | | | | 455 | | | | | 460 | | | | |
| Ser | Gly | Asn | Ile | Gly | Val | Phe | Asn | Val | Gly | Ser | Gly | Ser | Leu | Gly | Asn |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 |
| Tyr | Asn | Ile | Gly | Ser | Gly | Asn | Leu | Gly | Ile | Tyr | Asn | Ile | Gly | Phe | Gly |
| | | | | 485 | | | | | 490 | | | | | 495 | |
| Asn | Val | Gly | Asp | Tyr | Asn | Val | Gly | Phe | Gly | Asn | Ala | Gly | Asp | Phe | Asn |
| | | | 500 | | | | | 505 | | | | | 510 | | |
| Gln | Gly | Phe | Ala | Asn | Thr | Gly | Asn | Asn | Asn | Ile | Gly | Phe | Ala | Asn | Thr |
| | | 515 | | | | | | 520 | | | | 525 | | | |
| Gly | Asn | Asn | Asn | Ile | Gly | Ile | Gly | Leu | Ser | Gly | Asp | Asn | Gln | Gln | Gly |
| | 530 | | | | | 535 | | | | | 540 | | | | |
| Phe | Asn | Ile | Ala | Ser | Gly | Trp | Asn | Ser | Gly | Thr | Gly | Asn | Ser | Gly | Leu |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 |
| Phe | Asn | Ser | Gly | Thr | Asn | Asn | Val | Gly | Ile | Phe | Asn | Ala | Gly | Thr | Gly |
| | | | | 565 | | | | | 570 | | | | | 575 | |

Asn Val Gly Ile Ala Asn Ser Gly Thr Gly Asn Trp Gly Ile Gly Asn
 580 585 590
 Pro Gly Thr Asp Asn Thr Gly Ile Leu Asn Ala Gly Ser Tyr Asn Thr
 595 600 605
 Gly Ile Leu Asn Ala Gly Asp Phe Asn Thr Gly Phe Tyr Asn Thr Gly
 610 615 620
 Ser Tyr Asn Thr Gly Gly Phe Asn Val Gly Asn Thr Asn Thr Gly Asn
 625 630 635 640
 Phe Asn Val Gly Asp Thr Asn Thr Gly Ser Tyr Asn Pro Gly Asp Thr
 645 650 655
 Asn Thr Gly Phe Phe Asn Pro Gly Asn Val Asn Thr Gly Ala Phe Asp
 660 665 670
 Thr Gly Asp Phe Asn Asn Gly Phe Leu Val Ala Gly Asp Asn Glu Gly
 675 680 685
 Gln Ile Ala Ile Asp Leu Ser Val Thr Thr Pro Phe Ile Pro Ile Asn
 690 695 700
 Glu Gln Met Val Ile Asp Val His Asn Val Met Thr Phe Gly Gly Asn
 705 710 715 720
 Met Ile Thr Val Thr Glu Ala Ser Thr Val Phe Pro Gln Thr Phe Tyr
 725 730 735
 Leu Ser Gly Leu Phe Phe Phe Gly Pro Val Asn Leu Ser Ala Ser Thr
 740 745 750
 Leu Thr Val Pro Thr Ile Thr Leu Thr Ile Gly Gly Pro Thr Val Thr
 755 760 765
 Val Pro Ile Ser Ile Val Gly Ala Leu Glu Ser Arg Thr Ile Thr Phe
 770 775 780
 Leu Lys Ile Asp Pro Ala Pro Gly Ile Gly Asn Ser Thr Thr Asn Pro
 785 790 795 800
 Ser Ser Gly Phe Phe Asn Ser Gly Thr Gly Gly Thr Ser Gly Phe Gln
 805 810 815
 Asn Val Gly Gly Gly Ser Ser Gly Val Trp Asn Ser Gly Leu Ser Ser
 820 825 830
 Ala Ile Gly Asn Ser Gly Phe Gln Asn Leu Gly Ser Leu Gln Ser Gly
 835 840 845
 Trp Ala Asn Leu Gly Asn Ser Val Ser Gly Phe Phe Asn Thr Ser Thr
 850 855 860
 Val Asn Leu Ser Thr Pro Ala Asn Val Ser Gly Leu Asn Asn Ile Gly
 865 870 875 880
 Thr Asn Leu Ser Gly Val Phe Arg Gly Pro Thr Gly Thr Ile Phe Asn
 885 890 895
 Ala Gly Leu Ala Asn Leu Gly Gln Leu Asn Ile Gly Ser Ala Ser Cys
 900 905 910

202

Arg Ile Arg His Glu Leu Asp Thr Val Ser Thr Ile Ile Ser Ala Phe
 915 920 925

Cys Gly Ser Ala Ser Asp Glu Ser Asn Pro Gly Ser Val Ser Glu
 930 935 940

(2) INFORMATION FOR SEQ ID NO:205:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 53 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:205:

GGATCCATAT GGGCCATCAT CATCATCATC ACCTGATCGA CATCATCGGG ACC 53

(2) INFORMATION FOR SEQ ID NO:206:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:206:

CCTGAATTCG GGCCTCGGTT GCGCCGGCCT CATCTTGAAC GA 42

(2) INFORMATION FOR SEQ ID NO:207:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:207:

GGATCCTGCA GGCTCGAAGC CACCGAGCGG T 31

(2) INFORMATION FOR SEQ ID NO:208:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:208:
CTCTGAATTC AGCGCTGGAA ATCGTCGCGA T 31
- (2) INFORMATION FOR SEQ ID NO:209:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:209:
GGATCCAGCG CTCAGATGAA GACCGATGCC GCT 33
- (2) INFORMATION FOR SEQ ID NO:210:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:210:
GGATATCTGC AGAATTCAGG TTAAAGCCC ATTTGCGA 38
- (2) INFORMATION FOR SEQ ID NO:211:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:211:
CCGCATGCGA GCCACGTGCC CACACCGCCC 39
- (2) INFORMATION FOR SEQ ID NO:212:
- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:212:

CTTCATGGAA TTCTCAGGCC GGTAGGCTCC GCTGCGG

37

(2) INFORMATION FOR SEQ ID NO:213:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7676 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:213:

| | |
|---|------|
| TGGCGAATGG GACGCGCCCT GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG TGGTTACGGC | 60 |
| CAGCGTGACC GCTACACTTG CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC | 120 |
| CTTCTCGCC ACCTTCGCGG GCTTTCGCGC TCAGCTCTA AATCGCGGCT TCCCTTTAGG | 180 |
| GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCCAAAAA CTTGATTAGG GTGATGGTTC | 240 |
| ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT TTGACCTTGG AGTCCACGTT | 300 |
| CTTTAATAGT GGACTCTTGT TCCAACTGG AACACACTC AACCTATCT CCGTCTATTC | 360 |
| TTTGATTTA TAAGGATTT TCCCATTTT GGCCTATTGG TTAAGAAATG AGCTGATTTA | 420 |
| ACAAAAATTT AACGCGAATT TTAACAAAT ATTAACGTTT ACATTTTCAG GTGGCACTTT | 480 |
| TGGGGGAAAT GTGCGCGGAA CCCCTATTTT TTTATTTTTC TAATACATT CAAATATGTA | 540 |
| TCCCTCTATG AATTAATTCT TAGAAAACT CATCGAGCT CAAATGAAAC TGCAATTTAT | 600 |
| TCATATCAGG ATTATCAATA CCTATTTTTT GAAAAAGCCG TTTCTGTAAT GAAGGAGAAA | 660 |
| ACTCACCGAG GCAGTTCCAT AGGATGGCAA GATCCTGATA TCGGTCTGCG ATTCCSACTC | 720 |
| GTCCAACATC AATACAACCT ATTAATTTCC CTTCTTCAA AATAGCTTA TCAAGTGAGA | 780 |
| AATCACCATG AATTGAGACT GAATCGGCTG AGAATGGCAA AAGTTTATGC ATTCTTTTCC | 840 |
| AGACTTGTTC AACAGGCCAG CCATTACGCT CGTCATCAA ATCAGTCCCA TCAACCAAC | 900 |
| CGTTATTCTT TGGTGATTGC GCTGAGCGA GACGAATAC GCGATCCCTG TTAAAAGGAC | 960 |
| AATTACAAAC AGGAATCGAA TGCAACCGGC GCGGGAACAC TGCCAGCGCA TCAACAATAT | 1020 |
| TTTCACTTA ATCAAGATAT TCTTCTAATA CTTGGAATGC TGTTTTCCCG GGGATCGCAG | 1080 |
| TGGTGAGTAA CCATGCATCA TCAGGAGTAC GGTAAAAATG CTTGATGCTC GGAAGAGGCA | 1140 |

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|------|
| TAAATTCGGT | CAGCCASTTT | AGTCTGACCA | TCTCATCTGT | AACATCATTG | GCAACGGCTAC | 1200 |
| CTTTGCCATG | TTTCAGAAAC | AACTCTGGCG | CATCGGGCTT | CCCATACAA | CGATAGATTG | 1260 |
| TCGCACCTGA | TTGCCCCACA | TTATCGCGAG | CCGATTTATA | CCCAATATAA | TCAGCATCCA | 1320 |
| TGTTGGAAAT | TAATCGCGGC | CTAGAGCAAG | ACGTTTCCCG | TTGAATATGG | CTCATAACAC | 1380 |
| CCCTTGATTT | ACTGTTTATG | TAAGCAGACA | GTTTTATTGT | TCATGACCAA | AATCCCTTAA | 1440 |
| CGTGAGTTTT | CGTTCCACTG | AGCCTCAGAC | CCCGTAGAAA | AGATCAAAGG | ATCTTCTTGA | 1500 |
| GATCCTTTTT | TTCTGGCGGT | AATCTGCTGC | TTGCAACAA | AAAAACACCC | GCTACCAGCG | 1560 |
| GTGGTTTSTT | TGCGCGATCA | AGAGCTACCA | ACTCTTTTTT | CGAGCGTAAC | TGGCTTCAGC | 1620 |
| AGAGGCGAGA | TACCAATATC | TGTCCTTCTA | GTGTAGCCGT | AGTTAGGCCA | CCACTTCAAG | 1680 |
| AACTCTGTAG | CACCGCCTAC | ATACCTCGCT | CTGCTAATCC | TGTTACCAGT | GGCTGCTGCC | 1740 |
| AGTGGCGATA | ACTCGTGTCT | TACCGGCTTG | GACTCAAGAC | GATAGTTACC | GGATAAGGCG | 1800 |
| CAGCGGTGGG | GCTGAACGGG | GGGTTCTGTC | ACACAGCCCA | GCTTGGAGCG | AACGACCTAC | 1860 |
| ACCGAACTGA | GATACCTACA | GCCTGAGCTA | TGAGAAAGCG | CCACGCTTCC | CGAAGGAGAG | 1920 |
| AAGGCGGACA | GGTATCCGGT | AAGCGGCAGG | GTGGGACAG | GAGAGCGCAC | GAGGGAGCTT | 1980 |
| CCAGGGGGAA | ACGCTTGGTA | TCTTATATAT | CCTGTGCGGT | TTGCGCACCT | CTGACTTGAAG | 2040 |
| CGTCGATTTT | TGTGATGCTC | GTCAAGGGGG | CGGAGCCTAT | GGAAAAACGC | CAGCAACCGG | 2100 |
| GCCTTTTATC | GGTTCCCTGG | CTTTGCTGG | CTTTTTGCTC | ACATGTTCTT | TCCTGCGTTA | 2160 |
| TCCCTGATTT | CTGTGGATAA | CCSTATTACC | GCCTTTGAGT | GAGCTGATAC | CCCTGCGCGC | 2220 |
| AGCGGAACGA | CCGAGCGCAG | CGAGTCACTG | AGCGGGGAAG | CGGAAGAGCG | CGTGATGCGG | 2280 |
| TATTTTCTCC | TTACGCATCT | GTGCGGTATT | TCACACCGCA | TATATGGTGC | ACTCTCAGTA | 2340 |
| CAATUTGCTC | TGATGCGGCA | TAGTTAAGCC | AGTATACACT | CCGCTATCCG | TACCTGACTG | 2400 |
| GGTCAATGCT | GCGCCCGGAC | ACCCGCCAAC | ACCCGCTGAC | GCGCCCTGAC | GGGCTTGTCT | 2460 |
| GCTCCCGGCA | TCCGCTTACA | GACAAGCTGT | GACCGTCTCC | GGGAGCTGCA | TGTGTCAGAG | 2520 |
| GTTTTGACCG | TCATCACCGA | AACGCGCGAG | GCAGCTGCGG | TAAAGCTCAT | CAGCGTGGTC | 2580 |
| GTGAAGCGAT | TCACAGATGT | CTGCTGTGTC | ATCCGCGTCC | AGCTCGTTGA | GTTTCTCCAG | 2640 |
| AAGCTTAAAT | GTCTGGCTTC | TGATAAAGCG | GCCCATGTTA | AGGGCGGTTT | TTTCTGTGTT | 2700 |
| GGTCACTGAT | GCCTCCGTGT | AAGCGGGGAT | TCTGTTCTAT | GGGGTAATGA | TACCGATGAA | 2760 |
| ACGAGAGAGG | ATGCTCACGA | TACGGGTTAC | TGATGATGAA | CATGCCCCGT | TACTGGAACG | 2820 |
| TTGTGAGGGT | AAACAACCTG | CGGTATGGAT | GCGGCGGGAC | CAGAGAAAAA | TCACTCAGGG | 2880 |
| TCAATGCCAG | CGCTTCGTTA | ATACAGATGT | AGGTGTTCCA | CAGGGTAGCC | ASCAGCATCC | 2940 |
| TGCGATGCGG | ATCCGGAACA | TAATGGTGCA | GGGCGCTGAC | TTCCCGGTTT | CCAGACTTTA | 3000 |

| | | | | | | |
|------------|------------|------------|------------|-------------|-------------|------|
| CGAAACACGG | AAACCGAAGA | CCATTCATGT | TGTTGCTCAG | GTGCGAGAGG | TTTTGCAGCA | 3060 |
| GCAGTGGCTT | CACGTTGGCT | CGCGTATCGG | TCATTCATTG | TGCTAACCAG | TAGGGCAACC | 3120 |
| CCGCCAGGCT | AGCCCGGTCC | TCAACGACAG | GAGCAGGATC | ATGCCACACC | GTGGGGCCCG | 3180 |
| CATGCCGGCG | ATAATGGGCT | GCTTCTCGCC | GAAACGTTTG | GTGGCGGGAC | CAGTGACGAA | 3240 |
| GGCTTGAGCG | AGGGCGTGCA | AGATTCGGAA | TACCGCAAGC | GACAGGCCGA | TCATCGTGCT | 3300 |
| GCTCCAGCGA | AAGCGGTGCT | CGCCGAAAT | GACCCAGAGC | GCTGCCGGCA | CGTGCTCTAC | 3360 |
| GAGTTGCATG | ATAAAGAAAG | CAGTCATAAG | TGCGCGGACG | ATAGTCATGC | CCCCCGCCCA | 3420 |
| CGGGAAGGAG | CTGACTGGGT | TGAAGGCTCT | CAAGGGCATC | GCTCGAGATC | CCGTTGCCCTA | 3480 |
| ATCAGTGAGC | TAACTTACAT | TAACTGCTTT | GCGCTCACTG | CCCGCTTTCC | AGTCGGGAAA | 3540 |
| CGTGCTGTGC | CAGCTGCATT | AATGAATGGG | CCAACGCGCG | GGGAGAGGCG | GTTTGCGTAT | 3600 |
| TGGGCGCCAG | CGTGCTTTTT | CTTTTCACCA | GTGAGACGGG | CAACAGCTGA | TTGCCCTTCA | 3660 |
| CCGCTTGCGC | CTGAGAGAGT | TGCAGCAAGC | GCTCCAGGCT | GCTTTGCCCG | AGCAGCGGAA | 3720 |
| AATCCTGTTF | GATGGTGCTT | AACGGCGGGA | TATAACATGA | GCTGTCTTCC | GTATCGTCTT | 3780 |
| ATCCCACTAC | CGAGATATCC | GCACCAACGC | GCAGCCCGGA | CTCGGTAAAG | GCGCGCATTC | 3840 |
| CGCCAGGCGC | CATCTGATCG | TTGGCABCCA | GCATCGCAGT | GGGAACGATG | CCCTCATTTCA | 3900 |
| GCATTTGCAT | GGTTTGTGGA | AAACCGGACA | TGGCACTCCA | GTGGCCTTCC | CGTTCCGCTA | 3960 |
| TGGGCTGAAT | TTGATTGCGA | GTGAGATATT | TATGCCAGCC | AGCCAGACGC | AGACGCGCCC | 4020 |
| AGACGGAAGT | TAATGGGCCC | GCTAACAGCG | CGATTTGCTG | GTGACCCAAAT | GCGACCAGAT | 4080 |
| GCTCCAGGCC | CAGTGGGCTA | CGGTCTTCAT | GGGAGAAAAT | AATACTGTTG | ATGGGTGTCT | 4140 |
| GGTCAGAGAC | ATCAAGAAAT | AACCCCGGAA | CATTAGTCCA | GGCAGCTTCC | ACAGCAATGG | 4200 |
| CATCTGGGTC | ATCCAGCGGA | TAGTTAATGA | TCAGCCCACT | GACGCGTTGC | GCGAGAAGAT | 4260 |
| TGTGCACCGC | CGCTTTACAG | GCTTCGACGC | CGCTTCGTTT | TACCATCGAC | ACCACCAAGC | 4320 |
| TGGCAACCGG | TTGATGGGCG | CGAGATTAA | TCGCGCGGAC | AATTTGCGAC | CGCGCGTGCA | 4380 |
| GGGCGAGACT | GGAGGTGGCA | ACGCCAATCA | GCAACGACTG | TTTGCCCGCC | AGTTGTTGTG | 4440 |
| CCAGCGGCTT | GGGAATGTAA | TTGAGCTCCG | CCATCGCCGC | TTCCACTTTT | TCCCSCGTTT | 4500 |
| TCGCAAAAC | GTGGCTGGGC | TGTTTACCA | CGCGGGAAAC | GGTCTGATAA | GAGACACCGG | 4560 |
| CATACTCTGC | GACATCGTAT | AACGTTACTG | GTTCACATT | CACCAACCTG | AATTGACTCT | 4620 |
| CTTCCGGGCG | CTATCATGCC | ATACCGCGAA | AGTTTTCGCG | CCATTCGATG | GTGTCCGGGA | 4680 |
| TCGCAAGCT | CTCCCTTATG | CGACTCCTGC | ATTAGGAAGC | AGCCCAATAG | TAGGTTGAGG | 4740 |
| CGGTTGAGCA | CGCCCGCCGC | AAGGAATGGT | GCATGCAAGG | AGATGGCGCC | CAACAGTCCC | 4800 |
| CGGCGCACGG | GGCTTCCAC | CATACCCAGC | CGAABACAAG | CGCTCATGAG | CCCGAAGTGG | 4860 |

| | | | | | | |
|------------|------------|------------|------------|-------------|------------|------|
| CGAGCCCGAT | CTTCCCCATC | GGTGATGTGG | GCGATATAGG | CGCCAGCAAC | CGCACCTGTG | 4920 |
| GCGCCGGTGA | TGCGGGCCAC | GATGGGTCCG | GCGTAGAGCA | TGGAGATCTC | GATCCCCGGA | 4980 |
| AATTAAATCG | ACTCACTATA | GGGGAATTGT | GAGCGGATTA | CAATTCCCCCT | CTAGAAATTA | 5040 |
| TTTTGTTTTA | CTTTAAGAAG | GAGATATACA | TATGGGCCAT | CATCATCATC | ATCAAGTGAT | 5100 |
| CGACATCATC | GGCACCAGCC | CCACATCTTG | GGAACAGGCG | GCGGCGGAGG | CGGTCCAGCG | 5160 |
| GGCGCGGGAT | AGCGTCGATG | ACATCCGCGT | CGCTCGGCTC | ATTGAGCAGG | ACATGGCGGT | 5220 |
| GGACAGCGCC | GGCAAGATCA | CCTACCGCAT | CAAGCTCGAA | GTGTCTTTCA | AGATGAGGCC | 5280 |
| GGCGCAACCG | AGGGGCTCGA | AAGCACCGAG | CGGTTCCGCT | GAACCGGGCG | CGGGGCGCGG | 5340 |
| TACTGTCCGG | ACTACGCGCG | CGTCGTCCCG | GCTGACCTTG | GCGGAGACCG | GTAGCAGCGT | 5400 |
| GCTGTACCGG | CTGTTCAACC | TGTGGGGTCC | GGCTTTTCAC | GAGAGGATTC | CGAACCTCAC | 5460 |
| GATCACCGCT | CAGGGCAGCG | GTCTGTGTGC | CGGGATCCCG | CAGGCCGCGG | CGGGACCGGT | 5520 |
| CAACATTTGG | GGTTCGGAGG | GCTATCTGTC | GGAGGGTGAT | ATGGCCGCGC | ACAAGCGGCT | 5580 |
| GATGAACATC | GCGCTAGCCA | TCTCGGCTCA | GCAGGTCAAC | TACAACCTGC | CGGGAGTGAG | 5640 |
| CGAGCACCTC | AGGCTGAAGG | GAAAAGTCCT | GGCGGCCATG | TACCAGGGCA | CCATCAAAAC | 5700 |
| CTGGGACGAC | CGCAGATCG | CTGCGCTCAA | CGCGGGCTTG | AACCTGCGCG | GCACCGCGGT | 5760 |
| AGTTCCGCTG | CACCGGTCCG | ACGGGTCCGG | TGACACCTTC | TTGTTACCCG | AGTACCTGTC | 5820 |
| CAAGCAAGAT | CCCGAGGGCT | GGGGCAGATC | GCCCGGCTTC | GGCACCACCG | TGACTTTCCC | 5880 |
| GGCGGTGCGG | GGTGGGCTGG | GTGAGAACGG | CAACGCGGGC | ATGGTGCDCG | GTTCGCGCGA | 5940 |
| GACACCGGCG | TGCGTGGCCT | ATATCGGCAT | CAGTTTCCTC | GACCAGGCCA | GTCACCGGGG | 6000 |
| ACTCGGCGAG | GCCCAACTAG | GCAATAGCTC | TGGCAATTTC | TTGTTGCGCG | ACCGCGAAAG | 6060 |
| CATTCAAGCC | GCGGCGGGTG | GCTTCGCTAT | GAAAACCCCG | GCGAACCAGG | CGAATTTCAT | 6120 |
| GATCGACCGG | CCCGCGCGCG | ACGGCTACCC | GATCATTCAC | TACGAGTAGG | CCATCGTCAA | 6180 |
| CAACCGGCGA | AAGGAGCGCG | CCACCGCGCA | GACCTTGCAG | GCATTTCTGC | ACTGGGCGAT | 6240 |
| CACCGACCGC | AACAAGGCTT | CGTTCTCTGA | CCAGGTTTAT | TTCCAGCGCG | TGCCCGCGCG | 6300 |
| GGTGGTGAAG | TTGTCTGAGG | CGTTGATCGC | GACGATTTCC | AGCGCTGAGA | TGAAGACCGA | 6360 |
| TGCCGCTACC | CTCGCGCAGG | AGGCAGGTAA | TTTCGAGCGG | ATCTCCGCGG | ACCTGAAAAC | 6420 |
| CCAGATCGAC | CAGGTGGAT | CGACGCGAGG | TTGCTTTGAG | GGCAGTGGC | GCGGCGCGGC | 6480 |
| GGGACCGGCG | GCCAGGCGCG | CGGTGGTGCG | CTTCCAAGAA | GCAGCCAAAT | AGCAGAGCGA | 6540 |
| GGAACTCGAC | GAGATCTCGA | CGAATATTGG | TCAGGCGCGC | GTCCAATACT | CGAGGGCCGA | 6600 |
| CGAGGAGCAG | CAGCAGGCGC | TGTCTCTGCA | ATGGGCTTTT | GTGCGCACAA | CGGCGGCGCT | 6660 |
| GCGCGCGCTG | ACCGCTGCAG | CGCCACCGCG | ACCGGCGACA | CTGTGTGCGG | CGCCCGCGCG | 6720 |


```

GGCCGCCGCC AACACGCCGA ATGCCAGGCC GGGCGATCCC AACSCAGCAC CTCGCCGGGC 6780
CGACCCGAAC GCACCGCCGC CACCTGTCTT TCCCCAAAC GCACCCCAAC CTCTCCGGAT 6840
CGACAACCCG GTTGGAGGAT TCAGCTTCGC GTTGCCTGCT GCTTGGGTGG AGTCTGACGC 6900
CGCCCACTTC GACTACGGTT CAGCACTCCT CAGCAAAACC ACCGGGGACC CGCCATTTC 6960
CGGACAGCCG CGGCGGCTGG CCAATGACAC CCGTATCCTG CTCGGCCGGC TAGACCAAAA 7020
GCTTTACGCC AGCGCCGAAG CTACTGACTC CAAGGCCGCG GCGCGGTTGG GCTCGGACAT 7080
GGGTGAGTTC TATATGCCCT ACCCGGSCAC CCGGATCAAC CAGGAAACCG TCTCGCTTGA 7140
CGCCACCGGG GTGTCTGGAA GCGCGTCTTA TTACGAGTC AACTTCAGCG ATCCGAGTAA 7200
GCGGAACGGC CAGATCTGCA CGGGCTAAT CCGCTCGGCC GCGCGGACG CACCGGACGC 7260
CGGGCCCCCT CAGCGCTGGT TTGTGGTATG GCTCGGGACC GCGAACAACC CGGTGGACAA 7320
GGGCGCGGCC AAGGCGCTGG CCGAATCGAT CCGGCCTTTG GTCGCCCGGC CGCGGGGGCC 7380
GGCACCGGCT CTTGCAGAGC CCGCTCGGCT GCGGGCGCGG GCGGCGGAAG TCGCTCTTAC 7440
CCCGACGACA CCGACACCGC AGCGGACCTT ACCGGCTTGA GAATTCCTCA GATATCCATC 7500
ACACTGGGCG CCGCTCGAGC ACCACCACCA CCACCACTGA CATCCGGCTG CTAACAAAGC 7560
CGGAAAGSAA GCTGAGTTGG CTGCTGCCAC CCGTGAGCAA TAACTAGCAT AACCCCTTGG 7620
GGCTCTTAAA CGGCTCTTGA GGGTTTTTTT GGTGAAAGGA GGAATATAT CGGAT 7676

```

(2) INFORMATION FOR SEQ ID NO:214:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 802 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:

```

Met Gly His His His His His Val Ile Asp Ile Ile Gly Thr Ser
1           5           10           15
Pro Thr Ser Trp Glu Gln Ala Ala Ala Glu Ala Val Gln Arg Ala Arg
20           25           30
Asp Ser Val Asp Asp Ile Arg Val Ala Arg Val Ile Glu Gln Asp Met
35           40           45
Ala Val Asp Ser Ala Gly Lys Ile Thr Tyr Arg Ile Lys Leu Glu Val
50           55           60
Ser Phe Lys Met Arg Pro Ala Gln Pro Arg Gly Ser Lys Pro Pro Ser
65           70           75           80

```

209

Gly Ser Pro Glu Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro
 85 90 95
 Ala Ser Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr
 100 105 110
 Pro Leu Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn
 115 120 125
 Val Thr Ile Thr Ala Gln Gly Thr Gly Ser Gly Ala Gly Ile Ala Gln
 130 135 140
 Ala Ala Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser
 145 150 155 160
 Glu Gly Asp Met Ala Ala His Lys Gly Leu Met Asn Ile Ala Leu Ala
 165 170 175
 Ile Ser Ala Gln Gln Val Asn Tyr Asn Leu Pro Gly Val Ser Gln His
 180 185 190
 Leu Lys Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Gln Gly Thr Ile
 195 200 205
 Lys Thr Trp Asp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gly Val Asn
 210 215 220
 Leu Pro Gly Thr Ala Val Val Pro Leu His Arg Ser Asp Gly Ser Gly
 225 230 235 240
 Asp Thr Phe Leu Phe Thr Gln Tyr Leu Ser Lys Gln Asp Pro Gln Gly
 245 250 255
 Trp Gly Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val
 260 265 270
 Pro Gly Ala Leu Gly Glu Asn Gly Asn Gly Gly Met Val Thr Gly Cys
 275 280 285
 Ala Gln Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp
 290 295 300
 Gln Ala Ser Gln Arg Gly Leu Gly Glu Ala Gln Leu Gly Asn Ser Ser
 305 310 315 320
 Gly Asn Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala
 325 330 335
 Gly Phe Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp
 340 345 350
 Gly Pro Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile
 355 360 365
 Val Asn Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala
 370 375 380
 Phe Leu His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp
 385 390 395 400
 Gln Val His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp
 405 410 415

210

Ala Leu Ile Ala Thr Ile Ser Ser Ala Glu Met Lys Thr Asp Ala Ala
 420 425 430
 Thr Leu Ala Glu Glu Ala Gly Asn Phe Glu Arg Ile Ser Gly Asp Leu
 435 440 445
 Lys Thr Glu Ile Asp Glu Val Glu Ser Thr Ala Gly Ser Leu Glu Gly
 450 455 460
 Glu Trp Arg Gly Ala Ala Gly Thr Ala Ala Glu Ala Ala Val Val Arg
 465 470 475 480
 Phe Glu Glu Ala Ala Asn Lys Glu Lys Glu Glu Leu Asp Glu Ile Ser
 485 490 495
 Thr Asn Ile Arg Glu Ala Gly Val Glu Tyr Ser Arg Ala Asp Glu Glu
 500 505 510
 Glu Glu Glu Ala Leu Ser Ser Glu Met Gly Phe Val Pro Thr Thr Ala
 515 520 525
 Ala Ser Pro Pro Ser Thr Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro
 530 535 540
 Val Ala Pro Pro Pro Pro Ala Ala Ala Asn Thr Pro Asn Ala Glu Pro
 545 550 555 560
 Gly Asp Pro Asn Ala Ala Pro Pro Pro Ala Asp Pro Asn Ala Pro Pro
 565 570 575
 Pro Pro Val Ile Ala Pro Asn Ala Pro Glu Pro Val Arg Ile Asp Asn
 580 585 590
 Pro Val Gly Gly Phe Ser Phe Ala Leu Pro Ala Gly Trp Val Glu Ser
 595 600 605
 Asp Ala Ala His Phe Asp Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr
 610 615 620
 Gly Asp Pro Pro Phe Pro Gly Glu Pro Pro Pro Val Ala Asn Asp Thr
 625 630 635 640
 Arg Ile Val Leu Gly Arg Leu Asp Glu Lys Leu Tyr Ala Ser Ala Glu
 645 650 655
 Ala Thr Asp Ser Lys Ala Ala Ala Arg Leu Gly Ser Asp Met Gly Glu
 660 665 670
 Phe Tyr Met Pro Tyr Pro Gly Thr Arg Ile Asn Glu Glu Thr Val Ser
 675 680 685
 Leu Asp Ala Asn Gly Val Ser Gly Ser Ala Ser Tyr Tyr Glu Val Lys
 690 695 700
 Phe Ser Asp Pro Ser Lys Pro Asn Gly Glu Ile Trp Thr Gly Val Ile
 705 710 715 720
 Gly Ser Pro Ala Ala Asn Ala Pro Asp Ala Gly Pro Pro Glu Arg Trp
 725 730 735
 Phe Val Val Trp Leu Gly Thr Ala Asn Asn Pro Val Asp Lys Gly Ala

211

| 740 | | | | | 745 | | | | | 750 | | | | | | |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Ala | Lys | Ala | Leu | Ala | Glu | Ser | Ile | Arg | Pro | Leu | Val | Ala | Pro | Pro | Pro | |
| 755 | | | | | 760 | | | | | 765 | | | | | | |
| Ala | Pro | Ala | Pro | Ala | Pro | Ala | Glu | Pro | Ala | Pro | Ala | Pro | Ala | Pro | Ala | |
| 770 | | | | | 775 | | | | | 780 | | | | | | |
| Gly | Glu | Val | Ala | Pro | Thr | Pro | Thr | Thr | Pro | Thr | Pro | Gln | Arg | Thr | Leu | |
| 785 | | | | | 790 | | | | | 795 | | | | | 800 | |
| Pro Ala | | | | | | | | | | | | | | | | |

CLAIMS

1. A polypeptide comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID No. 120)
- (b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser; (SEQ ID No. 121)
- (c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID No. 122)
- (d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID No. 123)
- (e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID No. 124)
- (f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID No. 125)
- (g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Thr-Ala-Ala-Ser-Pro-Pro-Ser; (SEQ ID No. 126)
- (h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID No. 127)
- (i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn; (SEQ ID No. 128) and
- (j) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID No. 136)

wherein Xaa may be any amino acid.

2. A polypeptide comprising an immunogenic portion of an *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative

substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID No. 129) and
- (b) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID No. 137), wherein Xaa may be any amino acid.

3. A polypeptide comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101 or a complement thereof under moderately stringent conditions.

4. A polypeptide comprising an immunogenic portion of a *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 26-51, 138, 139, 163-183 and 201, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 26-51, 138, 139, 163-183 and 201 or a complement thereof under moderately stringent conditions.

5. A DNA molecule comprising a nucleotide sequence encoding a polypeptide according to any one of claims 1-4.

6. An expression vector comprising a DNA molecule according to claim 5.

7. A host cell transformed with an expression vector according to claim 6.

8. The host cell of claim 7 wherein the host cell is selected from the group consisting of *E. coli*, yeast and mammalian cells.

9. A pharmaceutical composition comprising one or more polypeptides according to any one of claims 1-4 and a physiologically acceptable carrier.

10. A pharmaceutical composition comprising one or more DNA molecules according to claim 5 and a physiologically acceptable carrier.

11. A pharmaceutical composition comprising one or more DNA sequences recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and a physiologically acceptable carrier.

12. A vaccine comprising one or more polypeptides according to any one of claims 1-4 and a non-specific immune response enhancer.

13. A vaccine comprising:
a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and
a non-specific immune response enhancer.

14. A vaccine comprising:
one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140 and 141, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and
a non-specific immune response enhancer.

15. The vaccine of claims 12-14 wherein the non-specific immune response enhancer is an adjuvant.

16. A vaccine comprising one or more DNA molecules according to claim 5 and a non-specific immune response enhancer.

17. A vaccine comprising one or more DNA sequences recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and a non-specific immune response enhancer.

18. The vaccine of claims 16 or 17 wherein the non-specific immune response enhancer is an adjuvant.

19. A pharmaceutical composition according to any one of claims 9-11, for use in the manufacture of a medicament for inducing protective immunity in a patient.

20. A vaccine according to any one of claims 12-18, for use in the manufacture of a medicament for inducing protective immunity in a patient.

21. A fusion protein comprising two or more polypeptides according to any one of claims 1-4.

22. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and ESAT-6.

23. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and the *M. tuberculosis* antigen 38 kD (SEQ ID NO:155).

24. A pharmaceutical composition comprising a fusion protein according to any one of claims 21-23 and a physiologically acceptable carrier.

25. A vaccine comprising a fusion protein according to any one of claims 21-23 and a non-specific immune response enhancer.

26. The vaccine of claim 25 wherein the non-specific immune response enhancer is an adjuvant.

27. A pharmaceutical composition according to claim 24, for use in the manufacture of a medicament for inducing protective immunity in a patient.

28. A vaccine according to claims 25 or 26, for use in the manufacture of a medicament for inducing protective immunity in a patient.

29. A method for detecting tuberculosis in a patient, comprising:

(a) contacting dermal cells of a patient with one or more polypeptides according to any one of claims 1-4; and

(b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

30. A method for detecting tuberculosis in a patient, comprising:

(a) contacting dermal cells of a patient with a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and

(b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

31. A method for detecting tuberculosis in a patient, comprising:

(a) contacting dermal cells of a patient with one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203; and

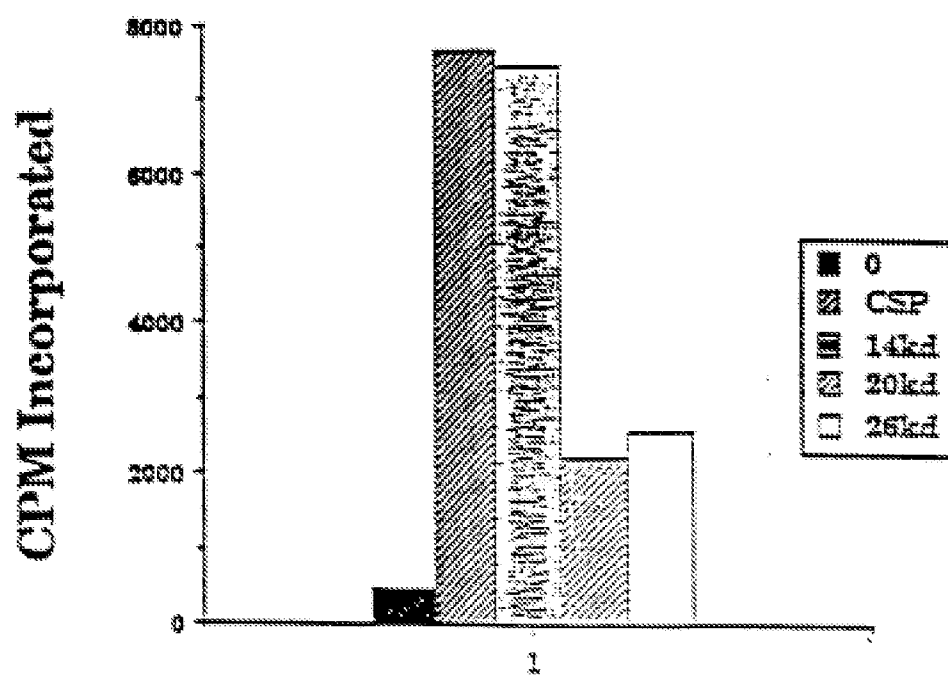
(b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

32. The method of any one of claims 29-31 wherein the immune response is induration.

33. A diagnostic kit comprising:
- (a) a polypeptide according to any one of claims 1-4; and
 - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
34. A diagnostic kit comprising:
- (a) a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and
 - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
35. A diagnostic kit comprising:
- (a) a polypeptide encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203; and
 - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
36. A diagnostic kit comprising:
- (a) a fusion protein according to any one of claims 21-23; and
 - (b) apparatus sufficient to contact said fusion protein with the dermal cells of a patient.
37. A fusion protein according to claim 23 comprising an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NO: 153 and 209.

1/11

D7 T Cell Proliferation



D7 IFN_g

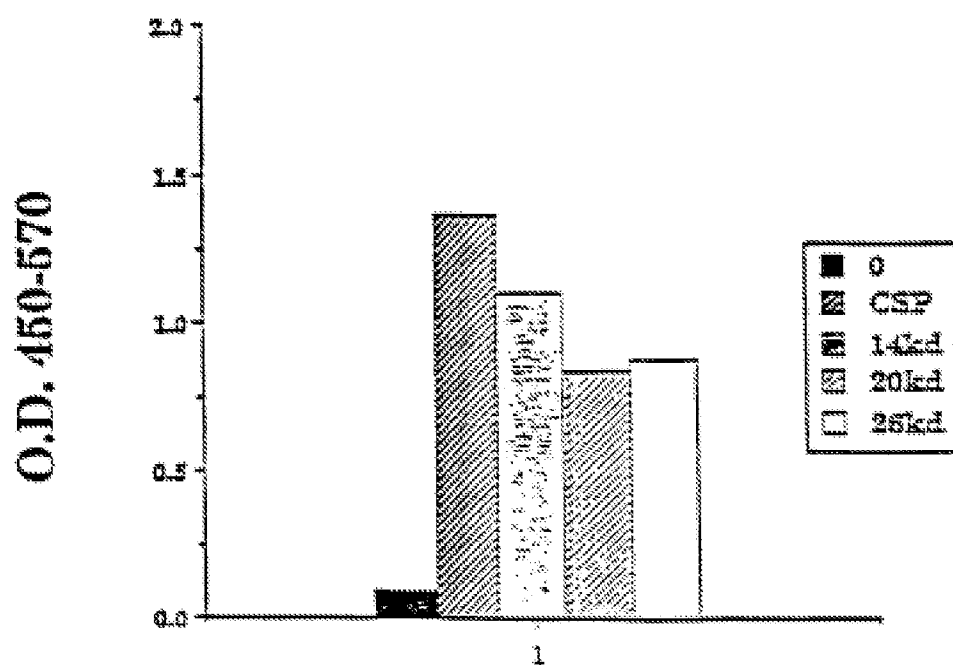
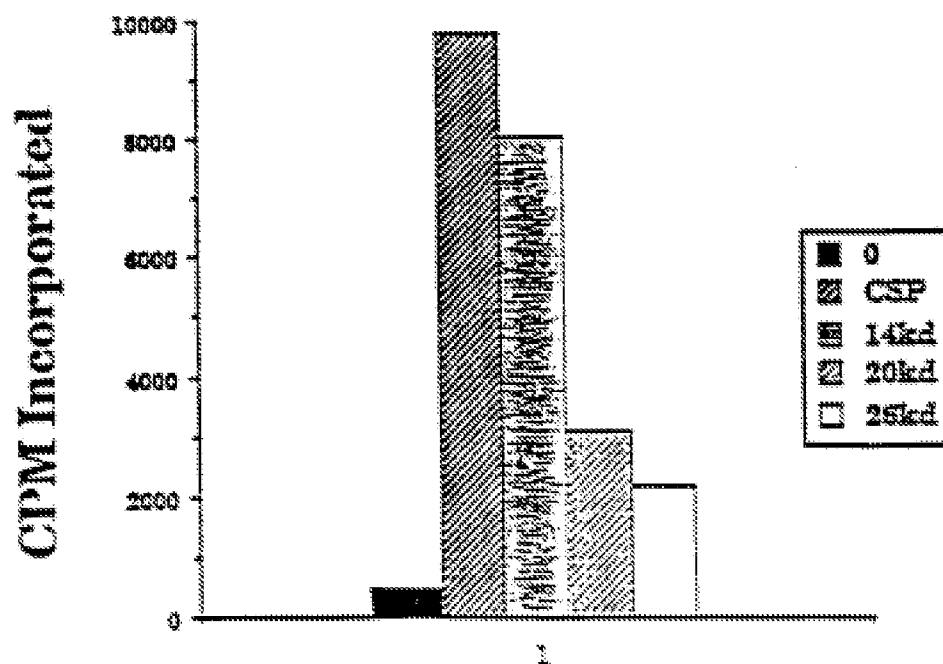


FIG. 1A

2/11

D160 T Cell Proliferation



D160 IFN γ

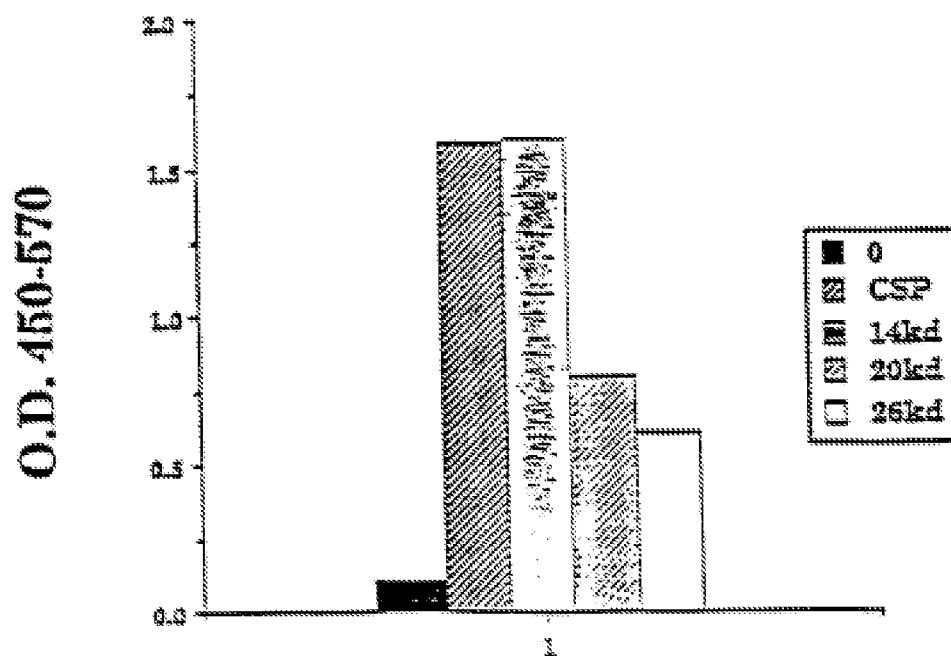


FIG. 1B

3/11

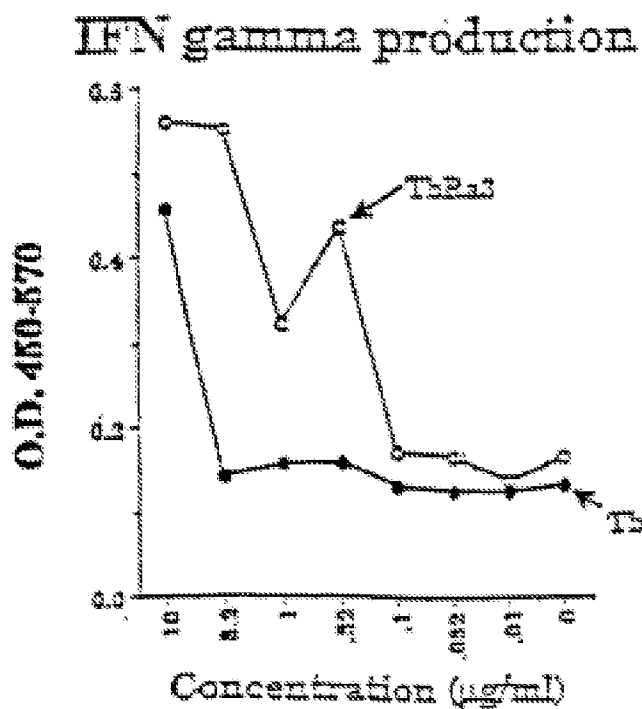
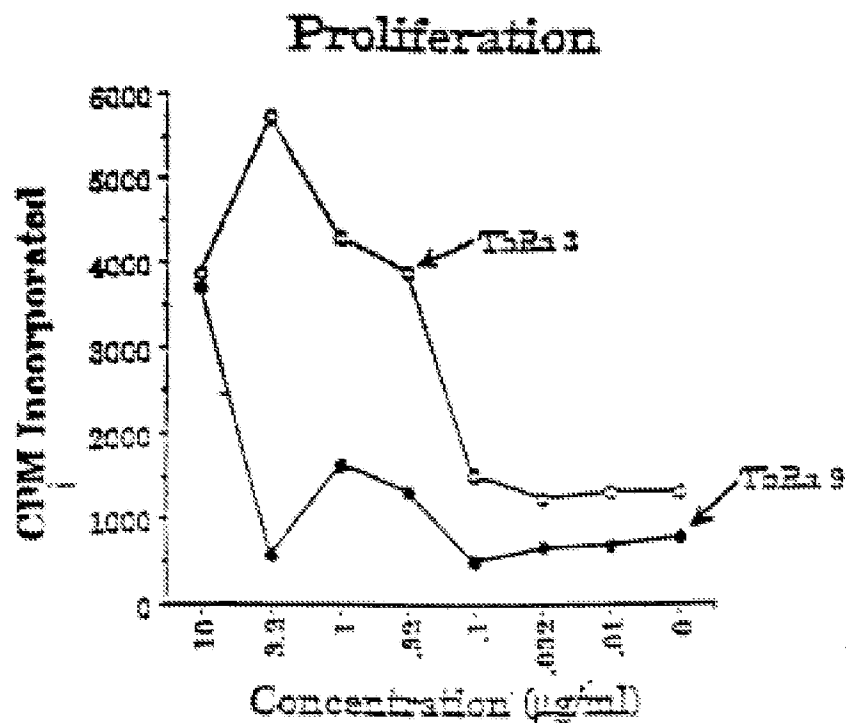
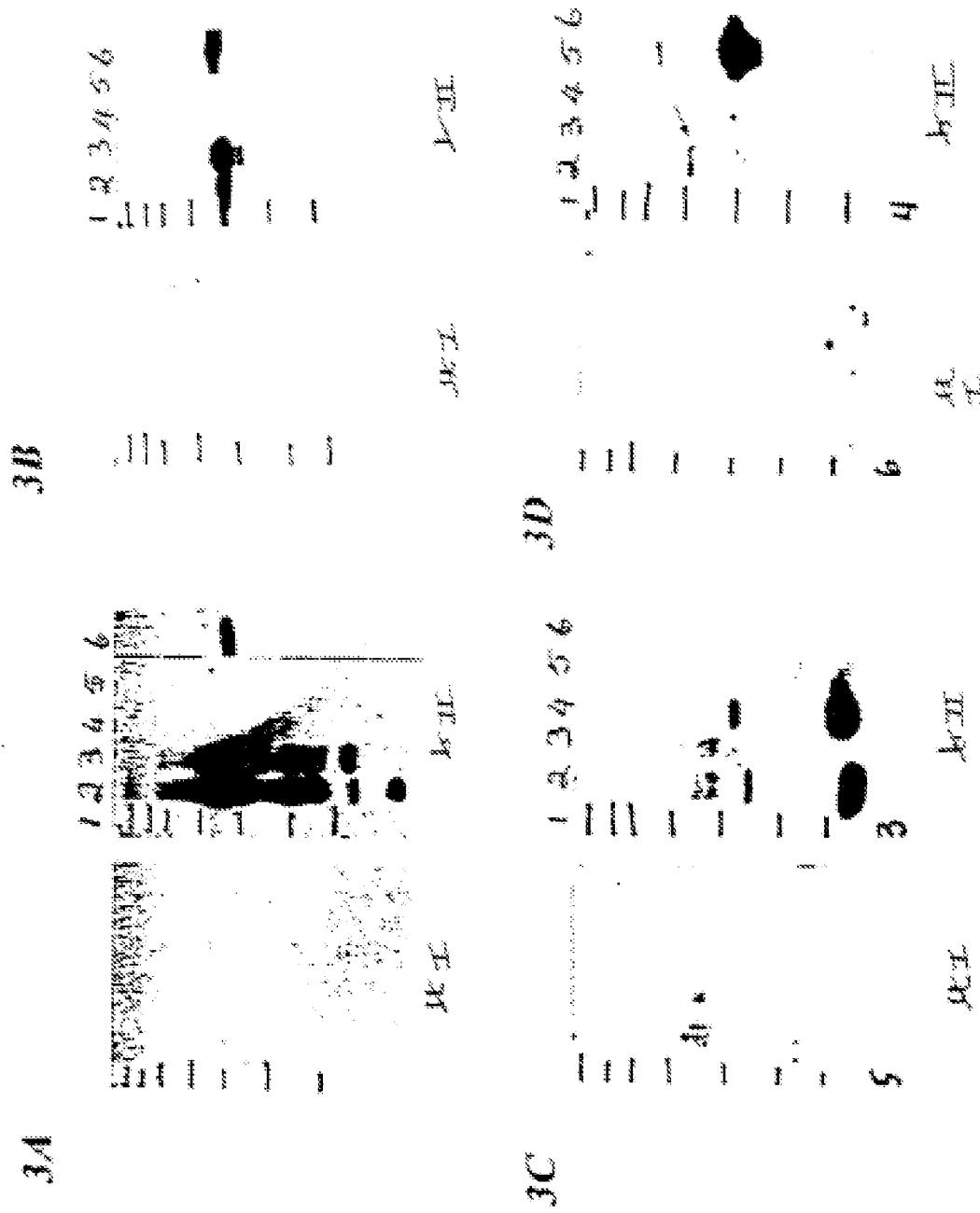


FIG. 2

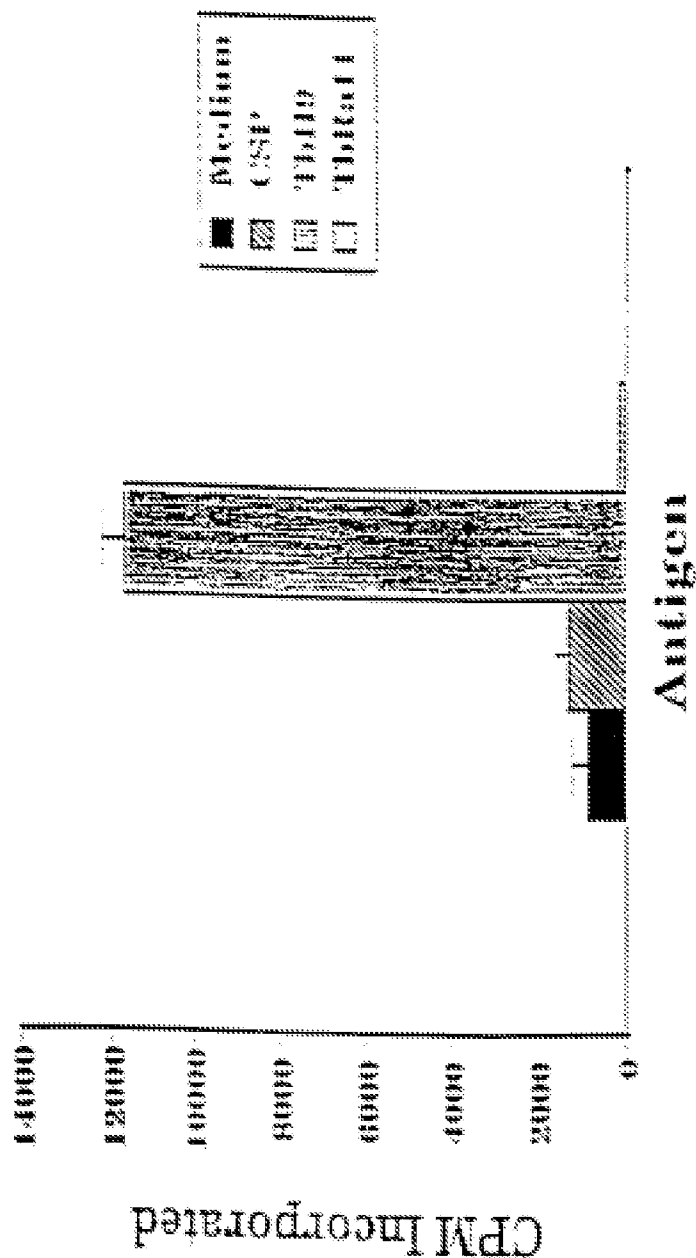
4/11



FIGS. 3 A-D

5/11

T cell clone 131TbH9 responds poorly to CSP



6/11

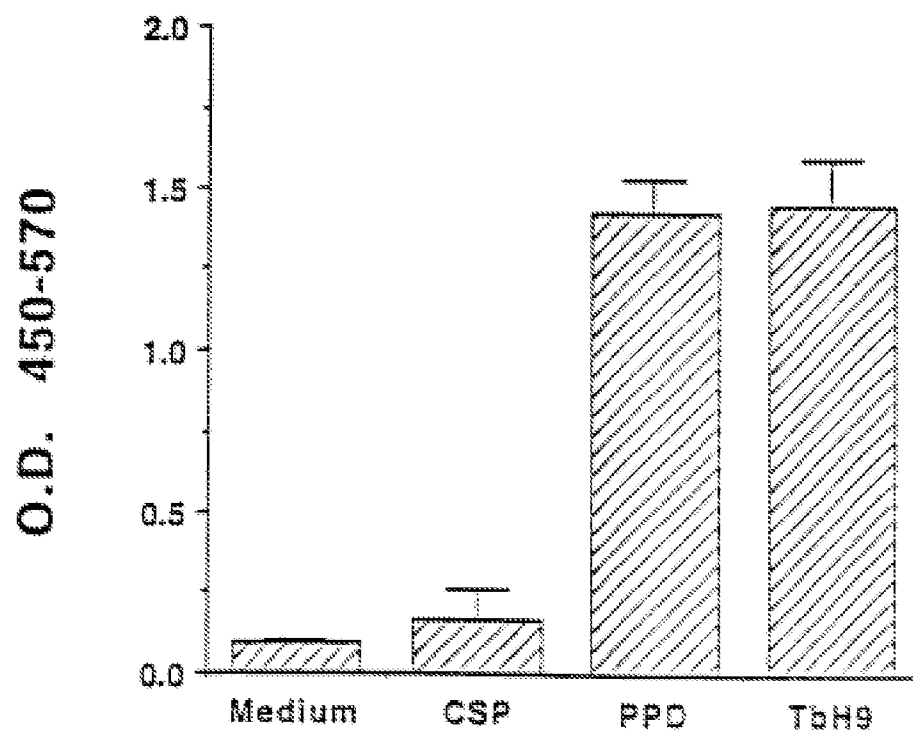
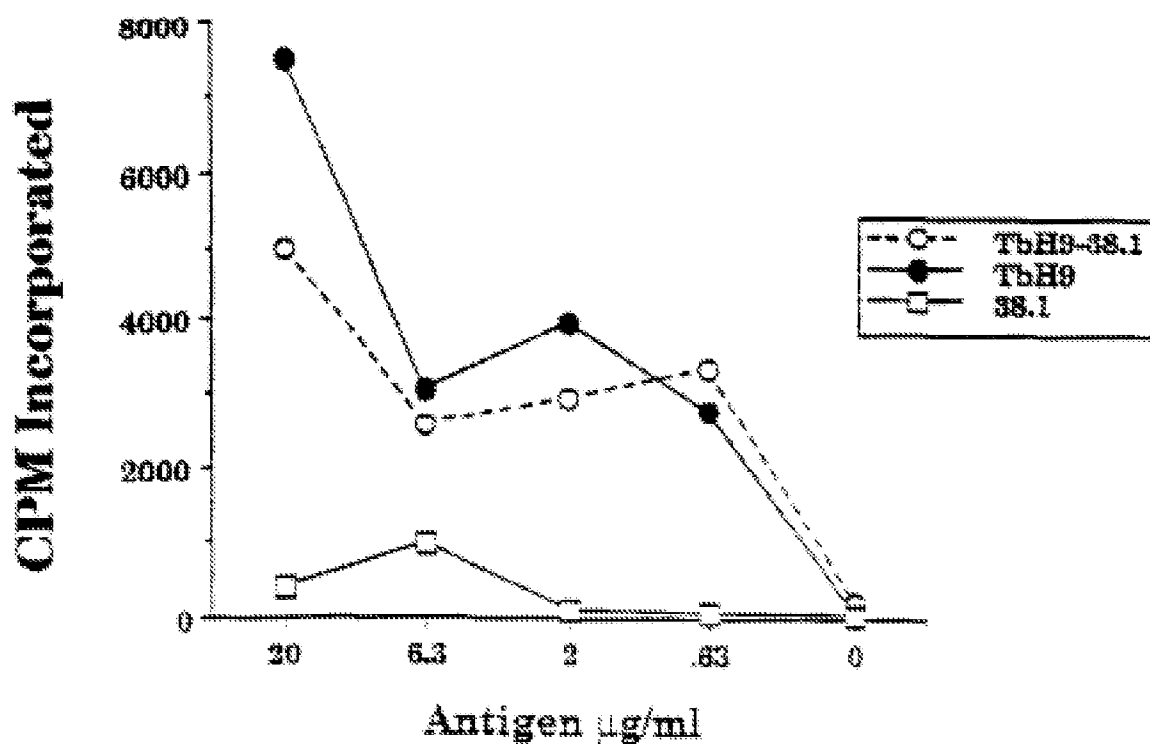
T Cell Clone PPD 800-10 IFN γ Production

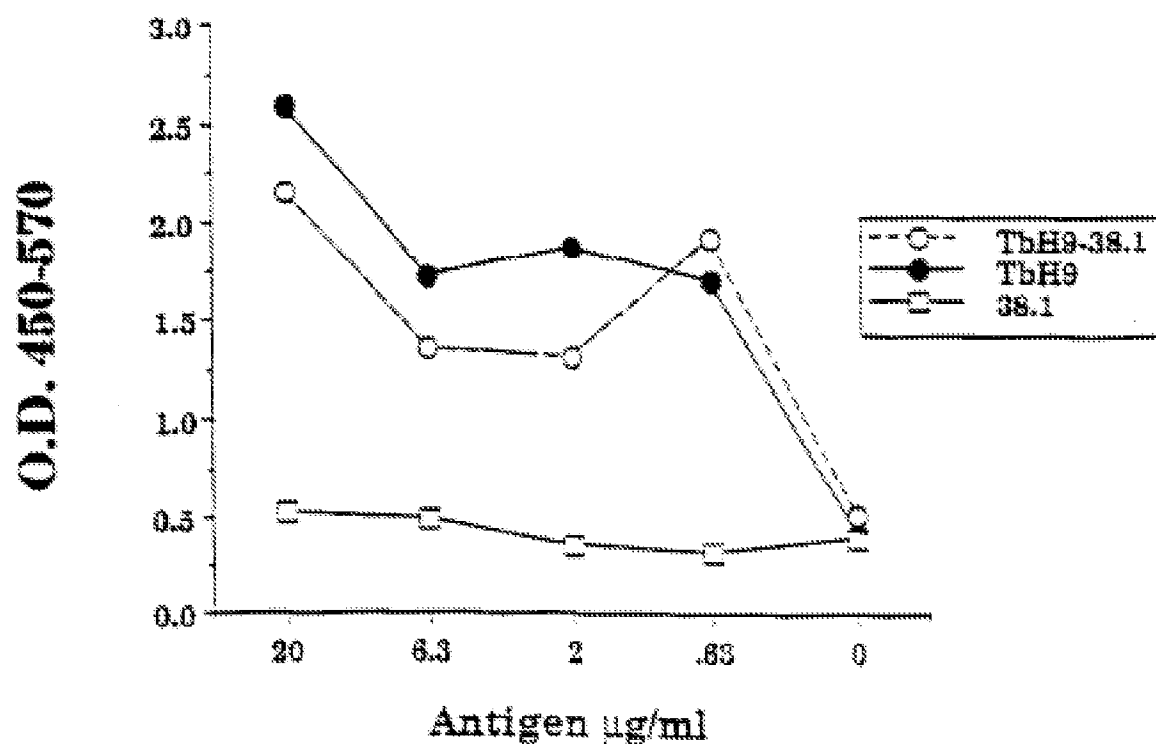
FIG. 4B

7/11

D131 T Cell Proliferation



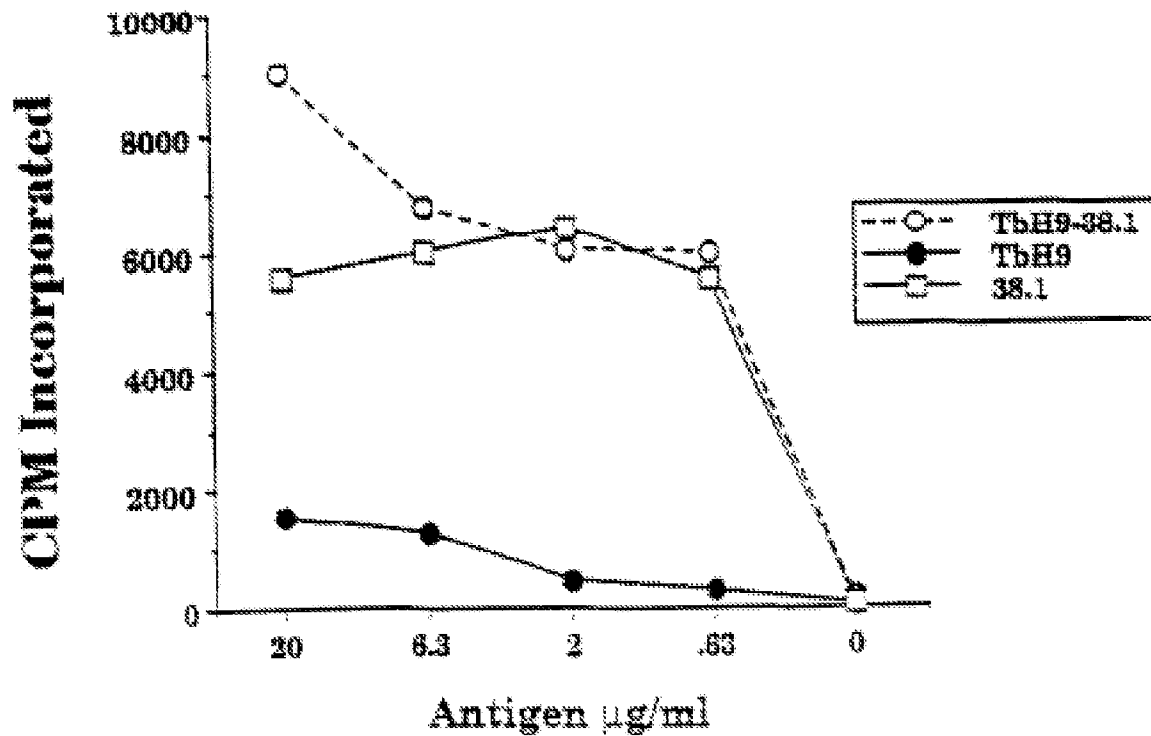
D131 IFNγ



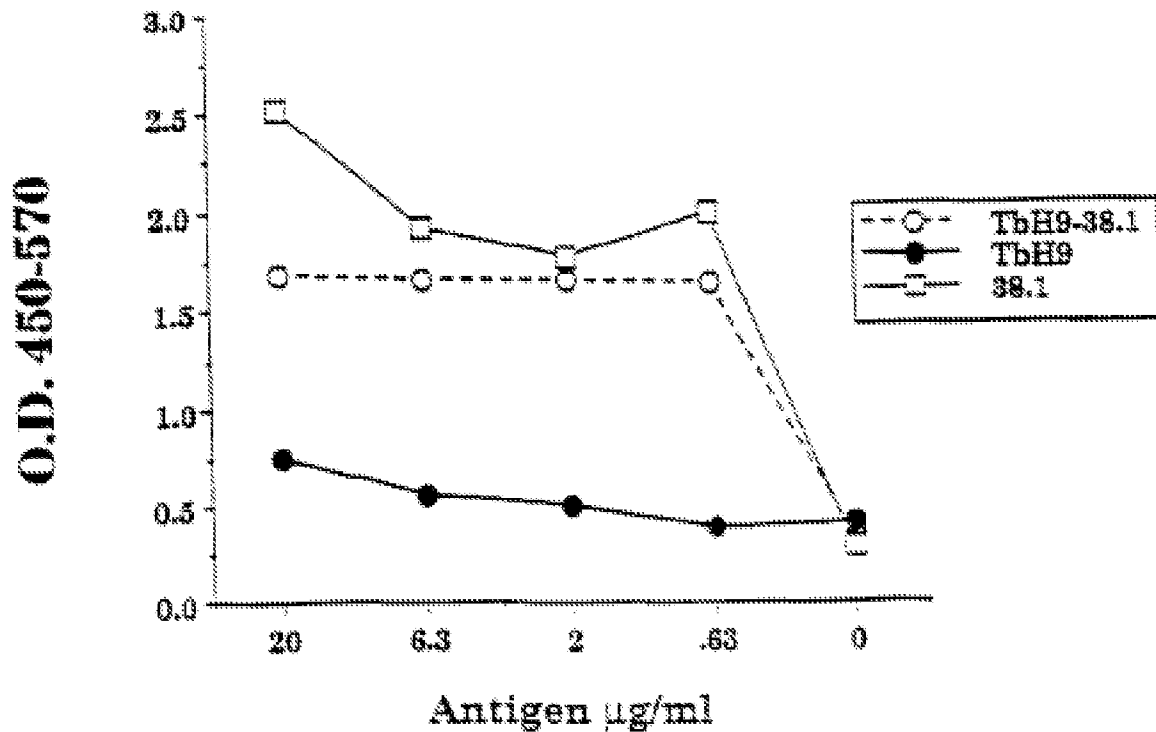
FIGS. 5 A-B

8/11

D184 T Cell Proliferation



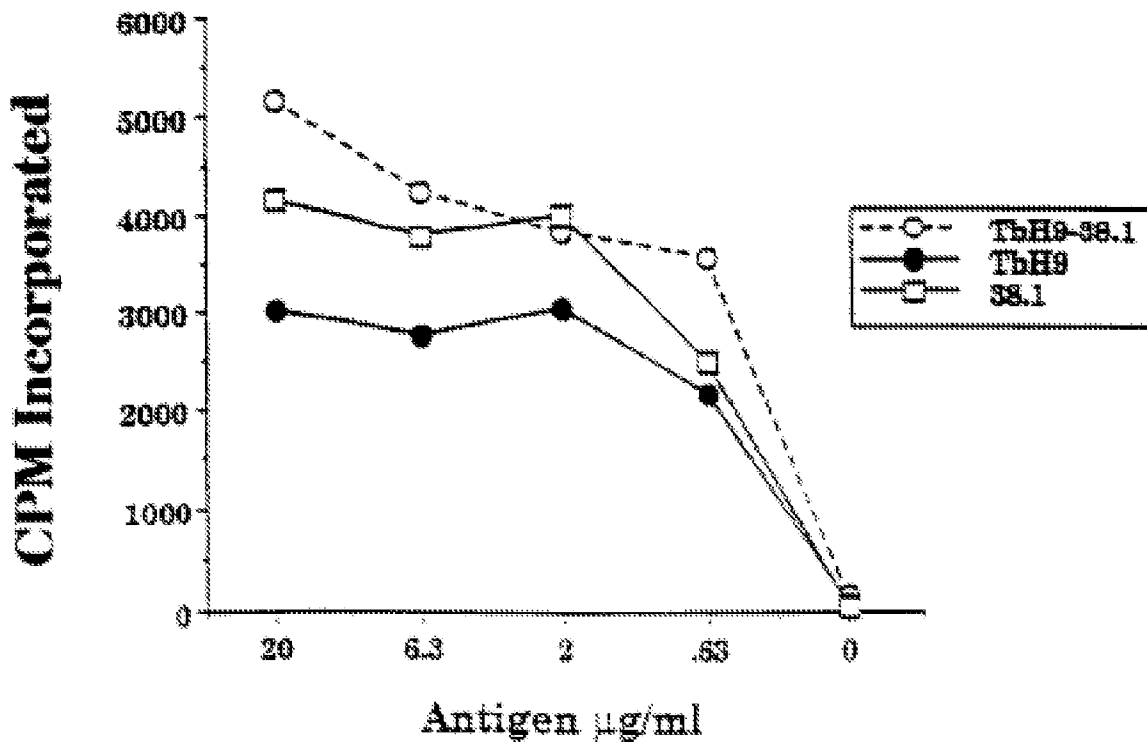
D184 IFN γ



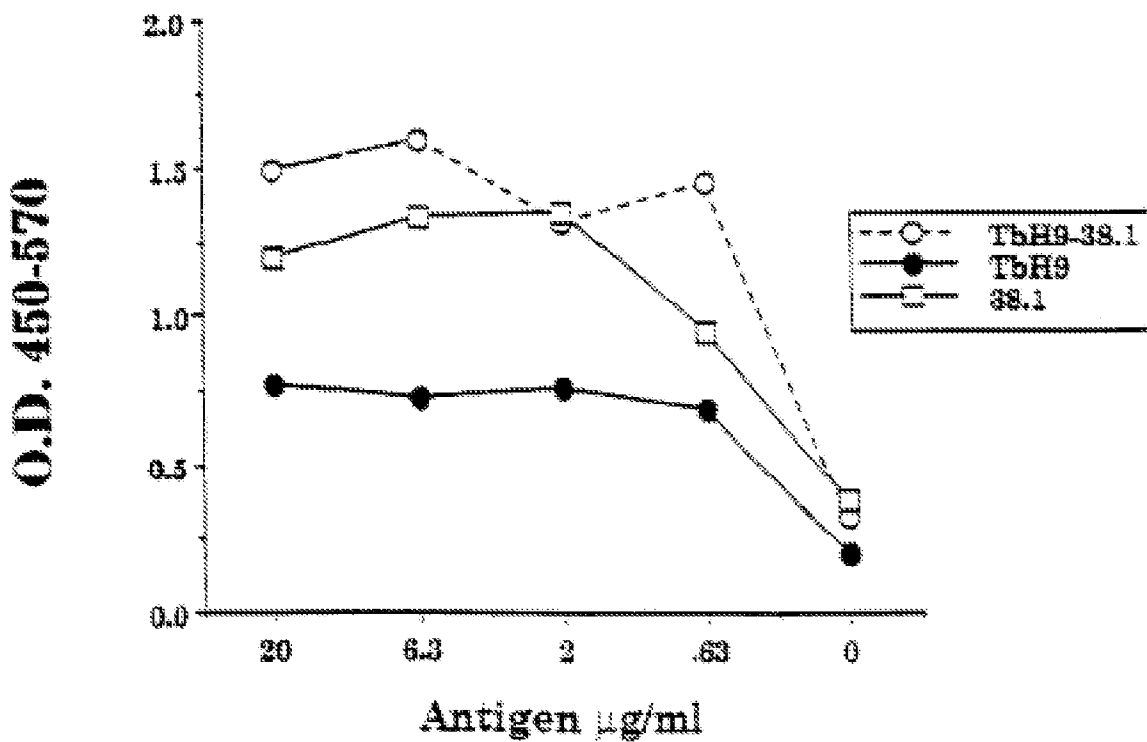
FIGS. 6 A-B

9/11

D201 T Cell Proliferation

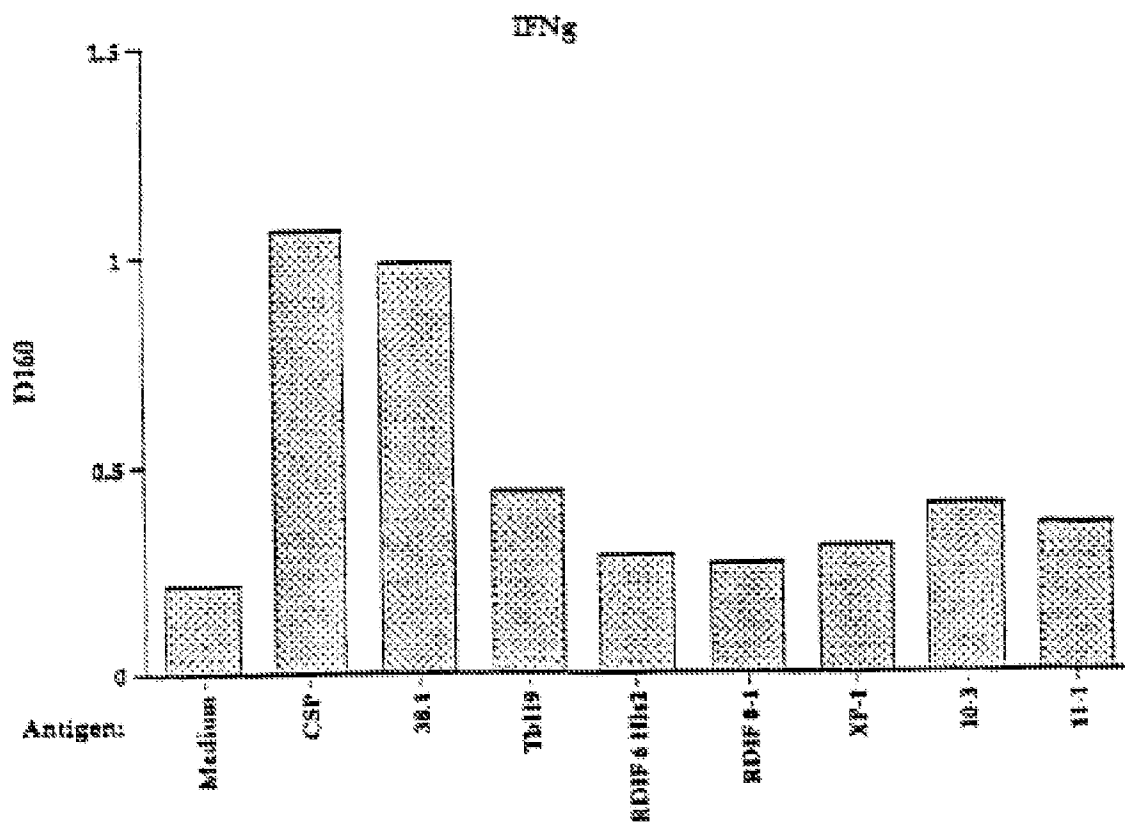
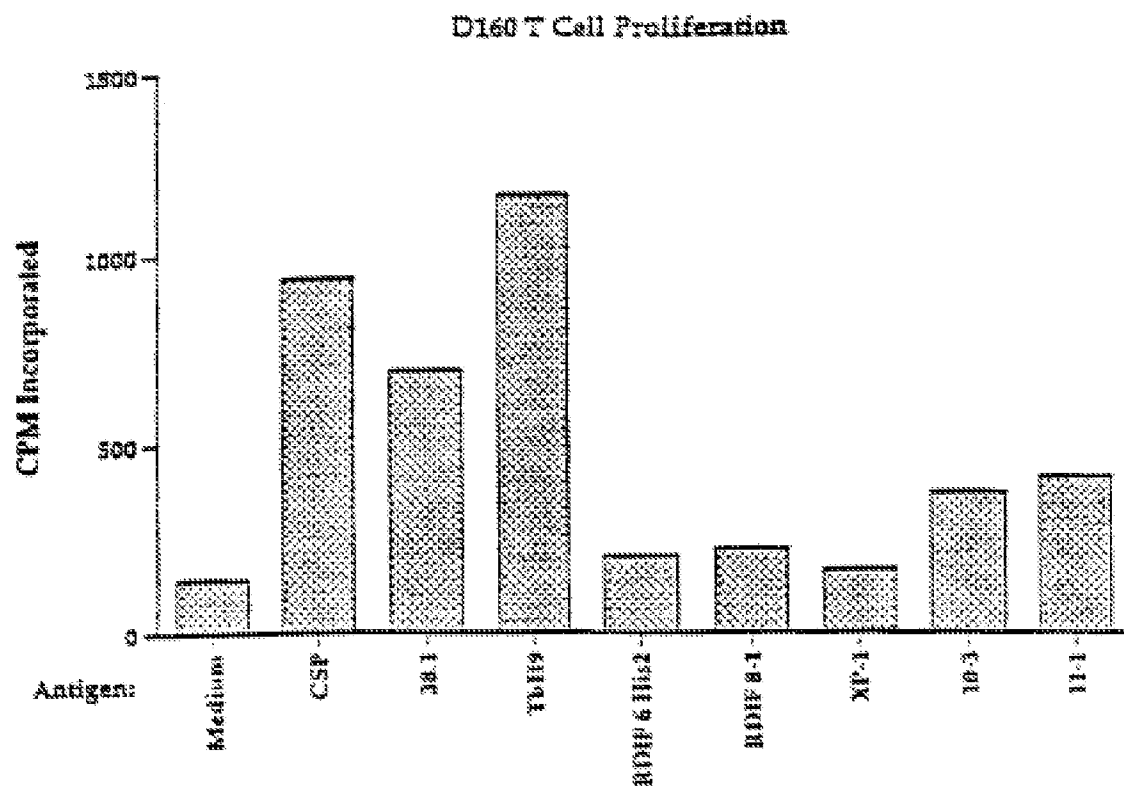


D201 IFN γ

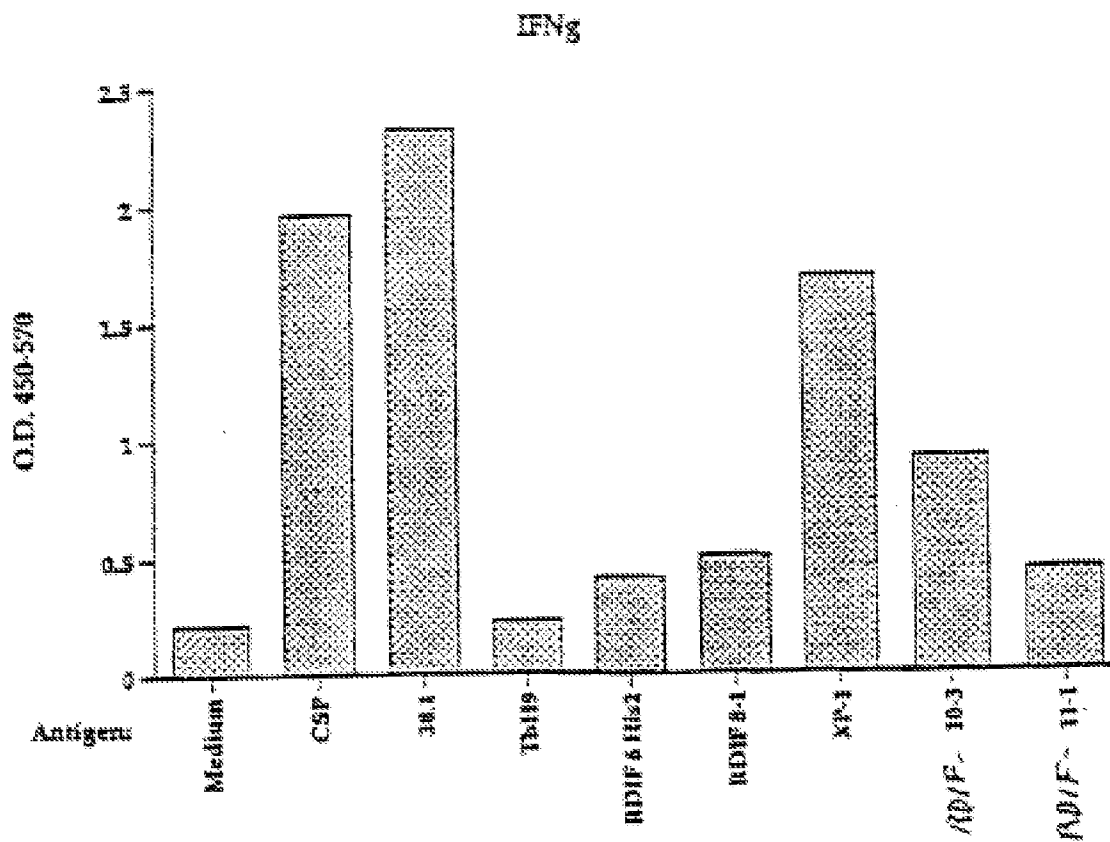
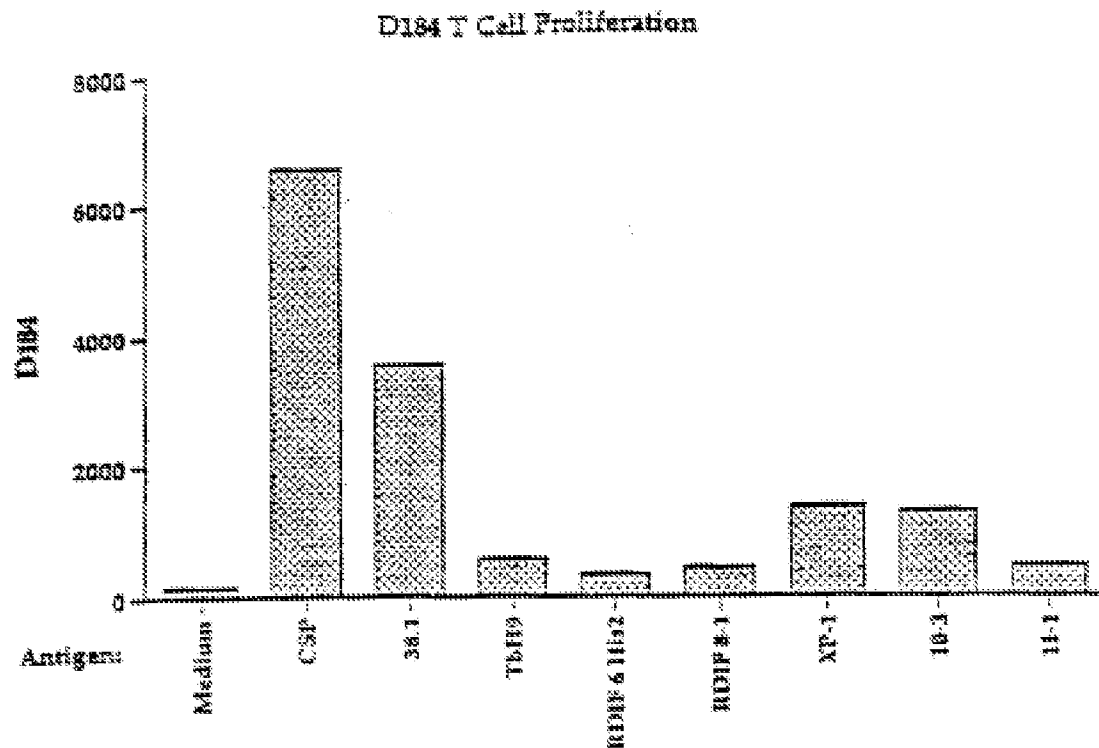


FIGS. 7 A-B

10/11

**FIGS. 8A-B**

11/11



FIGS. 9A-B